Volume 51 Numbers 25–26 2004 ISSN 0967-0645

SEVIER

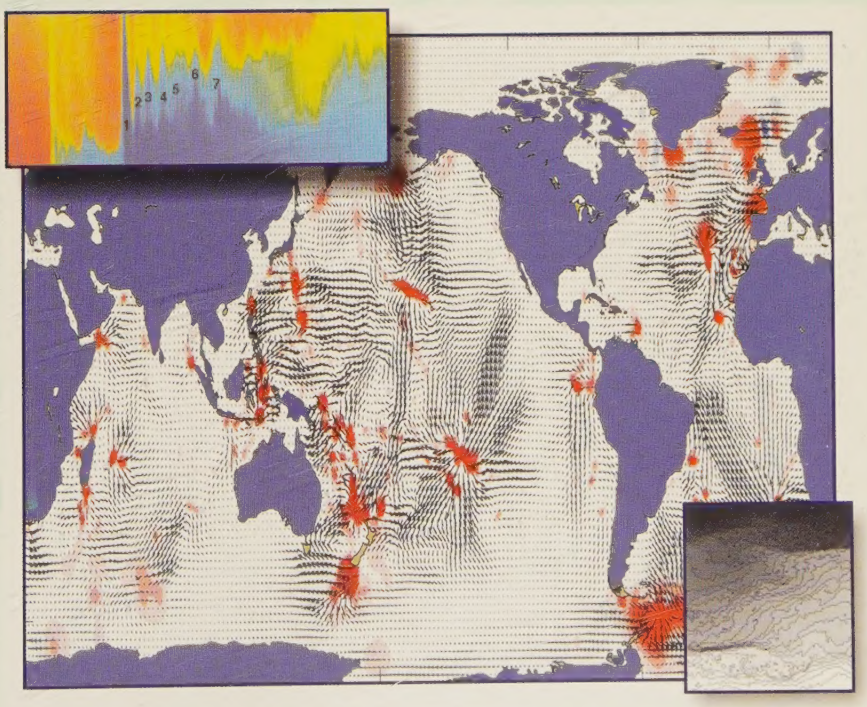
DEEP-SEA RESEARCH

PART II

n D. Milliman

Editors:
van Haren
St. Laurent
M. Marshall

Small and mesoscale processes and their impact on the large scale



Library
Pacific Islands Fisheries Science Center
Nat'l Marine Fisheries Service

Received Feb 2005

DEEP-SEA RESEARCH II

http://www.elsevier.com/locate/dsr2

Editor

PROFESSOR JOHN D. MILLIMAN: *Virginia Institute of Marine Science, School of Marine Science, The College of William and Mary, P.O. Box 1346, Gloucester Point, VA 23062–1346, USA*
E-mail: milliman@vims.edu

Associate Editors

J.W. MURRAY (*Chemistry*): *School of Oceanography, WB-10, University of Washington, Seattle, WA 98195, U.S.A. E-mail: jmurray@ocean.washington.edu*

FRIEDRICH A. SCHOTT: *Institut für Meereskunde an der Universität Kiel, Düsternbrooker Weg 20, D-24105 Kiel, Germany. E-mail: fschott@ifm.uni-kiel.de*

SHARON L. SMITH: *The Rosenstiel School, University of Miami, 4600 Rickenbacker Causeway, Miami, FL 33149, USA. E-mail: ssmith@rsmas.miami.edu*

Founder Editor: DR. MARY SEARS

Publishing Offices Elsevier Ltd, The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK. Tel: +44(0) 1865 843000; Fax: + 44(0) 1865 843010.

Advertising information. Advertising orders and enquiries can be sent to: ***USA, Canada and South America:*** Mr Tino DeCarlo, The Advertising Department, Elsevier Inc., 360 Park Avenue South, New York, NY 10010-1710, USA; phone: (+1) (212) 633 3815; fax: (+1) (212) 633 3820; e-mail: t.decarlo@elsevier.com. ***Europe and ROW:*** Katrina Barton, Advertising Department, Elsevier Ltd., 84 Theobald's Road, London WC1X 8RR, UK; phone (+44) 207 611 4117; fax: (+44) 207 611 4463; e-mail: k.barton@elsevier.com

USA mailing notice: *Deep-Sea Research II* (ISSN 0967-0645) is published annually in 26 issues by Elsevier Ltd. (P.O. Box 211, 1000 AE Amsterdam, The Netherlands). Annual subscription price in the USA US$ 3152 (valid in North, Central and South America), including air speed delivery.
USA POSTMASTER: Send address changes to *Deep-Sea Research II*, 655, Avenue of the Americas, New York, NY 10010.
AIRFREIGHT AND MAILING in the USA by Publications Expediting Inc., 200 Meacham Avenue, Elmont, NY 11003. Periodical postage rate paid at Jamaica, NY 11431.

Subscription Information

Published annually in 26 issues. Information on published and forthcoming issues can be found at the journal website.

Publication information: *Deep-Sea Research Part II* (ISSN 0967–0645). For 2004, volume 51 is scheduled for publication. Subscription prices are available upon request from the Publisher or from the Regional Sales Office nearest you or from this journal's website (http://www.elsevier.com./locate/dsr2). Further information is available on this journal and other Elsevier products through Elsevier's website: (http://www.elsevier.com). Subscriptions are accepted on a prepaid basis only and are entered on a calendar year basis. Issues are sent by standard mail (surface within Europe, air delivery outside Europe). Priority rates are available upon request. Claims for missing issues should be made within six months of the date of despatch.

Author enquiries. For enquiries relating to the submission of articles (including electronic submission), the preparation of electronic artwork, the status of accepted articles, author Frequently Asked Questions and any other enquiries relating to Elsevier, please consult http://authors.elsevier.com
Contact details for questions arising after acceptance of an article, especially those relating to proofs, are provided when an article is accepted for publication. Authors can also keep a track on the progress of their accepted article, and set up e-mail alerts informing them of changes to their manuscripts status, by using the "Track a paper" feature of Elsevier's Author Gateway (http: //authors.elsevier.com).

Orders, claims, and journal enquiries: please contact the Customer Service Department at the Regional Sales Office nearest you:
Orlando: Elsevier, Customer Service Department, 6277 Sea Harbor Drive, Orlando, FL 32887-4800, USA; phone: (+1) (877) 8397126 [toll free number for US customers], or (+1) (407) 3454020 [customers outside US]; fax: (+1) (407) 3631354; e-mail: usjcs@elsevier.com
Amsterdam: Elsevier, Customer Service Department, PO Box 211, 1000 AE Amsterdam, The Netherlands; phone: (+31) (20) 4853757; fax: (+31) (20) 4853432; e-mail: nlinfo-f@elsevier.com
Tokyo: Elsevier, Customer Service Department, 4F Higashi-Azabu, 1-Chome Bldg, 1-9-15 Higashi-Azabu, Minato-ku, Tokyo 106-0044, Japan; phone: (+81) (3) 5561 5037; fax: (+81) (3) 5561 5047; e-mail: jp.info@elsevier.com
Singapore: Elsevier, Customer Service Department, 3 Killiney Road, #08-01 Winsland House I, Singapore 239519; phone: (+65) 63490222; fax: (+65) 67331510; e-mail: asiainfo@elsevier.com

♾ The paper used in this publication meets the requirements of ANSI/NISO Z39. 48-1992 (Permanence of Paper).

© **2004 Elsevier Ltd. All rights reserved.**

Back Issues: Back issues of all previously published volumes are available direct from Elsevier Offices. Complete volumes and single issues can be purchased for 1999–2003. Earlier issues are available in high quality photo-duplicated copies as complete volumes only. Back volumes on microfilm are available from UMI, 300 North Zeeb Road, Ann Arbor, MI 48106-1346, USA.

Cover image: Global model on baroclinic energy flux and work (Fig. 13 in Simmons et al.) with inserts on small scale observations (left: near-bottom upslope propagating solibore; Fig. 11a in Hosegood et al.) and numerical simulation (right: mixing events in strong stratification; Fig. 6a in Bouruet-Aubertot et al.).

DEEP-SEA RESEARCH

PART II

Topical Studies in Oceanography

Small and mesoscale processes and their impact on the large scale

Guest Editors

Hans van Haren
NIOZ
1790 AB Den Burg
The Netherlands

Louis St. Laurent
Department of Oceanography
Florida State University
Tallahassee, FL 32306-4320, USA

David M. Marshall
Department of Meteorology
University of Reading
Reading RG6 6BB, UK

ELSEVIER

STATEMENT POLICY

Deep-Sea Research Part II: Topical Studies in Oceanography Publishes the topical issues arising from the ma international and interdisciplinary projects which are undertaken.

Shorter research papers are welcomed for *Deep-Sea Research Part I: Oceanographic Research Papers*. Acceptab topics for papers include:

(1) the results of original scientific research
(2) theoretical work of evident oceanographic applicability
(3) the solution of some instrumental or laboratory problem with evidence of its successful use, and
(4) applications of oceanography.

Papers dealing with research on continental shelves (water depths shallower than 100–200 m) in most cases should b submitted to our sister journal *Continental Shelf Research*.

Progress reports and review articles may be invited by the Editor but these are generally more appropriate fo *Progress in Oceanography*.

© **2004 Elsevier Ltd. All rights reserved.**

This journal and the individual contributions contained in it are protected under copyright by Elsevier Ltd, and the following term and conditions apply to their use:

Photocopying
Single photocopies of single articles may be made for personal use as allowed by national copyright laws. Permission of the publishe and payment of a fee is required for all other photocopying, including multiple or systematic copying, copying for advertising c promotional purposes, resale, and all forms of document delivery. Special rates are available for educational institutions that wish t make photocopies for non-profit educational classroom use.

Permissions may be sought directly from Elsevier Global Rights Department, P.O. Box 800, Oxford OX5 1DX, UK; phone: (+44 1865 843830, fax: (+44) 1865 853333, e-mail: permissionselsevier.co.uk. You may also contact Global Rights directly throug Elsevier's home page (http://www.elsevier.com), by selecting 'Obtaining Permissions'.

In the USA, users may clear permissions and make payments through the Copyright Clearance Center, Inc., 222 Rosewood Drive Danvers, MA 01923, USA; phone: (+1) (978) 7508400, fax: (+1) (978) 7504744, and in the UK through the Copyright Licensin Agency Rapid Clearance Service (CLARCS), 90 Tottenham Court Road, London W1P0LP, UK; phone: (+44) 20 7631 5555; fax (+44) 20 7631 5500. Other countries may have a local reprographic rights agency for payments.

Derivative Works
Subscribers may reproduce tables of contents or prepare lists of articles including abstracts for internal circulation within thei institutions. Permission of the publisher is required for resale or distribution outside the institution.

Permission of the publisher is required for all other derivative works, including compilations and translations.

Electronic Storage or Usage
Permission of the Publisher is required to store electronically any material contained in this journal, including any article or part of an article.

Except as outlined above, no part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior written permission of the publisher.

Address permissions requests to: Elsevier Global Rights Department, at the mail, fax and e-mail addresses noted above:

Notice
No responsibility is assumed by the Publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made.

Although all advertising material is expected to conform to ethical (medical) standards, inclusion in this publication does not constitute a guarantee or endorsement of the quality or value of such product or of the claims made of it by its manufacturer.

Small and mesoscale processes and their impact on the large scale

Volume 51 Nos. 25–26, 2004

CONTENTS

H. van Haren, L. St. Laurent and D. Marshall	2881	Preface
H. van Haren, L. St. Laurent and D. Marshall	2883	Small and mesoscale processes and their impact on the large scale: an introduction
B.R. Sutherland, M.R. Flynn and K. Dohan	2889	Internal wave excitation from a collapsing mixed region
R.H.J. Grimshaw and K.R. Khusnutdinova	2905	Internal waves in a three-layer bubbly waveguide
P. Bouruet-Aubertot, J. Sommeria, B. Le Cann and C.R. Koudella	2919	Intermittency of vertical density gradients at finescale and link with mixing processes
P. Hosegood and H. van Haren	2943	Near-bed solibores over the continental slope in the Faeroe-Shetland Channel
W.K. Dewar and D.A. Leonov	2973	Variability on steep, confined topography
T. Gerkema, F.-P.A. Lam and L.R.M. Maas	2995	Internal tides in the Bay of Biscay: conversion rates and seasonal effects
J. Isern-Fontanet, J. Font, E. García-Ladona, M. Emelianov, C. Millot and I. Taupier-Letage	3009	Spatial structure of anticyclonic eddies in the Algerian basin (Mediterranean Sea) analyzed using the Okubo–Weiss parameter
L.C. St. Laurent and J.D. Nash	3029	An examination of the radiative and dissipative properties of deep ocean internal tides
H.L. Simmons, R.W. Hallberg and B.K. Arbic	3043	Internal wave generation in a global baroclinic tide model
B.K. Arbic, S.T. Garner, R.W. Hallberg and H.L. Simmons	3069	The accuracy of surface elevations in forward global barotropic and baroclinic tide models

CONTENTS Direct

This journal is part of **ContentsDirect**, the *free* alerting service which sends tables of contents by e-mail for Elsevier books and journals. You can register for **ContentsDirect** online at: **www.elsevier.com/locate/contentsdirect**

ELSEVIER

Deep-Sea Research II 51 (2004) 2881

DEEP-SEA RESEARCH
PART II

www.elsevier.com/locate/dsr2

Preface

During the 2003 EGS–AGU–EUG Joint Assembly held in Nice, France, 06–11 April 2003, the turn-out was tremendous: "...first- and last-joint meeting of AGU, EUG and EGS ...and the more than 10,000 participants from over 90 countries made it one the largest gatherings of geoscientists ever" (EOS 84(17) 29 April 2003, p. 159). Some 14,000 abstracts were submitted. Of these, some 70 ended up in the Ocean Sciences session OS3 on "Small and mesoscale processes and their impact on the large scale". Within the session approximately 40% (27 contributions) were presented orally and none was cancelled. The discussion was lively, indicating that the subject studying detailed ocean processes ranging from internal waves to eddies is highly topical in present-day (physical) oceanographic research. Some 10 of the OS3-contributions eventually found their way as a paper to this special issue of Deep-Sea Research II under a tight scheme of deadlines. These papers reflect more or less the contributions presented at the Joint Assembly meeting, they were not preselected. This publication within 1 year after submission of the manuscripts could not have been established without the generous help of the anonymous referees. We hereby thank them for their support.

Hans van Haren
Royal Netherlands Institute for Sea Research (NIOZ), P.O. Box 59, 1790 AB Den Burg, The Netherlands
E-mail address: hansvh@nioz.nl

Louis St. Laurent
Department of Oceanography, Florida State University, Rm 407 OSB, West Call Street, 32306-4320 Tallahassee, FL, United States of America
E-mail address: lous@ocean.fsu.edu

David Marshall
Department of Meteorology, University of Reading, P.O. Box 243, Reading RG6 6BB, UK
E-mail address: davidm@met.rdg.ac.uk

0967-0645/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dsr2.2004.09.009

ELSEVIER

Deep-Sea Research II 51 (2004) 2883–2887

DEEP-SEA RESEARCH
PART II

www.elsevier.com/locate/dsr2

Small and mesoscale processes and their impact on the large scale: an introduction

Hans van Haren[a,*], Louis St. Laurent[b], David Marshall[c]

[a]*Royal Netherlands Institute for Sea Research (NIOZ), PO Box 59, 1790 AB Den Burg, The Netherlands*
[b]*Department of Oceanography, Florida State University, Tallahassee, FL, 32306-4320, USA*
[c]*Department of Meteorology, University of Reading, PO Box 243, Reading RG6 6BB, UK*

Abstract

Great progress has been made in understanding the role of small- and mesoscale processes on the large-scale structure and circulation of the oceans. However, many questions remain regarding the sensitivity of the large-scale ocean circulation and hence the earth's climate, to rates of stirring, mixing, and dissipation supported by ocean eddies, internal waves, and boundary layers. Although such relatively small-scale oceanic phenomena are ubiquitous, we lack detailed knowledge about their interaction with the large-scale motions, as well as their interactions amongst themselves.

Contributions are presented from observational, theoretical, and modeling studies that focus on small- and mesoscale ocean processes, and some parameterizations of these processes for use in studies of the large-scale circulation. The range of topics is relatively broad, with most focus on internal wave generation, flow–topography interaction, and small- and mesoscale processes involved in ocean mixing. General conclusions emphasize the notion of localized mixing and the importance of interaction between waves, eddies, and topography, and the importance of interactions between oceanic motions and the background density stratification.
© 2004 Elsevier Ltd. All rights reserved.

1. Introduction

Recent public concern over the impact of human activity on climate has sustained strong scientific interest in this subject. It has long been recognized that the large water content near the earth's surface, whether in vapour form as in the atmosphere, in liquid form as in the oceans, or in solid form as in ice, is crucial for understanding the variability of weather and climate. Each of the three states of water has received its own field of scientific interest, and many problems remain to be resolved. Since most of the water is found in the oceans, this has become the focus of many studies.

*Corresponding author. Tel.: 31 0 222 369451; fax: 31 0 222 319674

E-mail addresses: hansvh@nioz.nl (H. van Haren), lous@ocean.fsu.edu (L.S. Laurent), davidm@met.rdg.ac.uk (D. Marshall).

0967-0645/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dsr2.2004.09.010

Early studies of the role of the oceans in climate and climatic change focused on modelling the large-scale ocean circulation. However, it quickly became apparent that the large-scale circulation is not maintained solely by large-scale processes. Rather, small- and mesoscale processes, responsible for turbulent diffusion and the redistribution of heat, salt, and suspended materials, can have a very significant impact on the larger scale. While the importance of small-scale processes has long been recognized, it is difficult to study large-scale ocean circulation and small-scale processes at the same time. This has led to two distinct branches of physical oceanographic studies. One branch is focused on the proper implementation ('parameterization') of the net effects of small-scale processes on large-scale motions. The other branch is focused on the detailed dynamics of the relevant small-scale processes.

Recent work has broadened our understanding of the dominant small- and mesoscale processes in the oceans (e.g., see the reviews in Munk and Wunsch, 1998; Wunsch and Ferrari, 2004). Specifically, these studies have revealed the importance of tidal forcing and surface wind forcing on both the small- and large-scale dynamics of the ocean. Motions at near-inertial and tidal frequencies are communicated into the ocean interior by propagating internal waves that may break in certain areas, thereby causing mixing. Especially important is the recognition of possible mixing in localized areas instead of globally uniform and omnipresent mixing (Wunsch and Ferrari, 2004). The interaction of ocean oscillatory motions with bottom topography is not just important for mixing, but also for the generation of internal tides. Typical horizontal length scales for internal waves are O(10 km), while typical vertical length scales are O(100–1000 m). These scales and smaller are considered "small-scales" here. Bottom topography also appears to be important for the intensification of boundary currents that may become unstable (meander) and split-off eddies of "mesoscale" size, O(100 km) horizontally and O(1000 m) vertically. Such eddies can be long-lived (1–2 yr), but eventually dissipate, thereby distributing their contents in a different environment than where they were generated. However, for these small- and mesoscale processes, the mecha isms controlling the energy cascade to irreversib turbulent dissipation are still in question.

The papers in this volume represent a range studies working to reveal some of the aforemer tioned details for parameterization of small- an mesoscale processes in models of the large-sca circulation. These studies cover a wide range scales and use a wide variety of techniques fro laboratory and numerical process modelling, t large-scale modelling and analysis of observation

2. Small and mesoscale processes...

As noted above, one of the main candidates fo ocean mixing is the breaking of internal waves which are mainly generated externally via tides o atmospheric forcing. A quite different scenario i presented by Sutherland et al. (2004), who discuss the generation of internal gravity waves following a mixing event. By way of laboratory experiments and fully non-linear numerical simulations, they study the collapse of a mixed region of uniform density into a surrounding region of stratified fluid, the latter thereby able to support internal waves. Complex interactions are found, for specific types of stratification, when internal waves are excited with such large amplitudes that the collapsing mixed region is distorted through strong interactions with the waves. This study shows that complex interactions are possible between internal waves, allowing for breaking and subsequent "re"generation [possibly at different frequencies] in the intrinsic system of deep-ocean motions. A similar unexpected interaction is observed in a two-dimensional (2-D) model for the near-surface ocean, where air- bubbles influence the vertical stratification and consequently the propagation of internal gravity waves (Grimshaw and Khusnutdinova, 2004). This model shows that the effective stratification due to bubbles leads to the possibility of existence of internal [gravity] waves and a change of their dispersion relation in the otherwise homogeneous upper mixed layer.

The above studies emphasize the importance of the interaction between internal waves and their supporting background. This theme is found in

many other studies in this volume. At the finescale end, Bouruet-Aúbertot et al. (2004) study intermittent variations in density gradients caused by non-linear interactions among internal gravity waves. Sheets of strong stable density gradients result from convergent motion produced by surrounding breaking events. From their statistics of layers and sheets, the authors infer a vertical eddy diffusivity of $5 \times 10^{-4}\,m^2\,s^{-1}$. This value is five times larger than the canonical value considered necessary to maintain the large-scale meridional overturning circulation in the ocean (Munk, 1966) and an order of magnitude larger than typical values observed and attributed to internal wave breaking (e.g., Gregg and Kunze, 1991; van Haren et al., 1999). This and other studies lead to the suggestion (Garrett, 1990, and others) that particular topographic sites may be important for internal wave breaking and thus localized enhanced mixing.

Indeed very high vertical diffusivities of $O(10^{-1}\,m^2\,s^{-1})$ are observed in a detailed observational study using fast and accurate sampling instrumentation above a sloping side of the Faeroe-Shetland Channel (Hosegood and van Haren, 2004). Such high values are observed during a very limited period of time only, associated with turbulence generated by the passage of strongly non-linear wave trains (solibores) propagating up the slope. As the events are separated by approximately 4 days, the long-term average vertical diffusivity exceed $10^{-4}\,m^2\,s^{-1}$, despite the relatively strong tidal currents in the area. It is concluded that solibores probably cannot sustain sufficient deep-sea mixing required for the large-scale circulation. In contrast, daily averaged sediment fluxes measured near the bed were O(100) times larger than the background value during the day of one particular event, implying solibores are the dominant sediment transport mechanism over the slope, despite their intermittent character. The sediment fluxes (and hence the turbulence intensity) vary with the propagation direction of the solibore. The times between the events (larger than a tidal period) lead to the suggestion that boundary current variations or eddies govern the deep (~500–800 m) mixing events at the slope, possibly triggered by variations in passing atmospheric disturbances. This quasi-4-day sub-inertial periodicity does not seem to be influenced by the bottom topography, as the slope is very smooth in the area of investigation.

The effect of stratification intersecting sloping topography, specifically abrupt topography, on sub-inertial boundary waves is studied by Dewar and Leonov (2004) using analytical and numerical methods. Both baroclinic and barotropic waves are considered. It is argued that such waves possess a number of unusual linear dispersive and non-linear steepening tendencies. As a result, baroclinic waves on finite topographic slopes are shown to mix barotropic and baroclinic dynamics in a manner distinct from those on weak topographic slopes. This study confirms the recent notion of the importance of complex 'rough' topography of steep slopes and rapid along-slope variability (e.g., canyons). On the other hand, as demonstrated in the previous study, smooth sloping sides can support vigorous processes as well, and are still considered important for substantial internal wave generation.

Internal tides in one of the most renowned source areas, the shelf-break of the Bay of Biscay, are studied by Gerkema et al. (2004) using a numerical, linear, hydrostatic internal-tide generation model. They compare their model results with unique observations made near the continental shelf-break in early summer using a towed acoustic Doppler current profiler. The modelling focus is on the seasonal changes in the generation and dynamics. The authors find that in reproducing the observed patterns with the numerical model, the presence of the seasonal pycnocline is essential. However, the main region of generation for internal tides is below the seasonal pycnocline, and the permanent pycnocline determines the integrated energy conversion rates. Future observational investigations should seek to confirm the result, from the numerical model, that the energy radiated in internal wave beams becomes diffuse after a few reflections with the surface and bottom boundaries, even without considering dissipation. In addition, the reflection points could be important sites for mixing.

Localized enhanced mixing not only occurs near internal wave generation sites, but also in the open

ocean where mesoscale eddies are found. By definition, mesoscale eddies carry water masses that are different from their surroundings. Mixing eventually occurs as the eddies dissipate. Many studies have examined the characteristics of mesoscale eddies near the surface (e.g., Gulf Stream warm core rings) and in the ocean interior (e.g., Mediterranean outflow eddies—Meddies). Mesoscale eddies have also been studied using satellite observations, such as in the study by Isern-Fontanet et al. (2004). They examine three eddies observed in the western Mediterranean Sea using satellite altimetry and sateillite observations of sea-surface temperature, both referenced to in situ data. They find that the observed eddies have a spatial structure in close resemblance to that found in coherent vortices of 2-D turbulence. The distinction between the coherent vorticity-dominated core of the eddy and the outer region dominated by deformation is important for understanding the role of transport and stirring properties associated with the eddies.

2.1. ...And their impact on the large scale

The role of small-scale processes has also been assessed relative to the large-scale energy budgets of the ocean. St. Laurent and Nash (2004) present measurements of internal tide energy flux from the Mid-Atlantic Ridge and the Hawaiian Ridge. They find significant differences in both level and modal content in the radiated internal tide at these distinct generation sites. The lateral decay scale of energy flux generated at the Hawaiian Ridge is found to be O(1000) km, due to the stability of the low modes which characterize the radiation. In contrast, the energy flux generated at the Mid-Atlantic Ridge is comprised of higher modes which dissipate more efficiently, leading to a decay scale of O(100) km. Despite these differences, St. Laurent and Nash point out that the level of turbulent mixing at each site remains comparable.

Simmons et al. (2004) also consider the large-scale implications of the internal tide energy budget. They present numerical simulations of internal tides at 10-km lateral scales over the global domain. For the first time, patterns of baroclinic radiation are visible on basin scales, and complex patterns of energy flux generation a radiation emerge. Simmons et al. show that th energy flux estimates agree well with independe estimates for prominent generation sites. This i noteworthy accomplishment, as their model so tion results solely from the forward problem solving the governing baroclinic equations subje to the astronomical tidal forcing.

Arbic et al. (2004) also consider forwa solutions for the ocean tides, but strive to achie accuracy in the barotropic (surface) mode as primary goal. Their solution for the surface ti meets an accuracy previously met only by mode that use assimilation of observed data. To achie this, Arbic et al. use a topographic drag schem based on the energy transfer from the barotrop to baroclinic tides. Both one- and two-laye models are examined, and the inclusion of th barotropic to baroclinic transfer is shown to b crucial in providing accurate forward estimates o the tides. Thus, Arbic et al. have demonstrated clear link between a global-scale phenomena, th surface tides, and the small-scale process o internal tide generation.

3. Suggestions for future research

As is clear from the papers presented in this volume, we are still lacking so much knowledge o the detailed dynamics of the small- and mesoscale processes, that it remains hard to give conclusive closure schemes and parameterizations of their effects on the larger scale. In particular, the representation of small- and mesoscale processes in models of the global-scale climate remains as a major goal for future research. Recent studies on small- and mesoscale motions have shown that such parameterizations should be spatially variable, as generation, dissipation, and mixing occur in localized areas only. Topography, or more precisely, topography with certain specifications like matching the slope angle of internal wave rays, seems a key factor influencing the generation and dissipation of waves and eddies. More detailed studies are needed on the precise scale dependence of topography and on the types of interaction with oceanic motions. Furthermore, studies focused on

the interactions between internal waves, eddies, and their oceanic background, are still needed. This requires finescale modelling down to the smallest scales and resolution of the relevant timescales of variation. It also requires new observational techniques resolving finescales in time of order 1 s, and in space of order 1 m vertically and of order 100 m horizontally, with high accuracy for sufficiently long periods to examine climate variations of interest. Given the advances in modern computing, such goals are becoming more and more realistic.

References

Arbic, B.K., Garner, S.T., Hallberg, R.W., Simmons, H.L., 2004. The uses and accuracy of a forward global tide model: effects of topographic drag, self-attraction and loading, and stratification. Deep-Sea Research II, this issue [doi:10.1016/j.dsr2.2004.09.014].

Bouruet-Aubertot, P., Sommeria, J., Le Cann, B., Koudella, C.R., 2004. Intermittency of vertical density gradients at finescale and link with mixing processes. Deep-Sea Research II, this issue [doi:10.1016/j.dsr2.2004.09.007].

Dewar, W., Leonov, D.A., 2004. Variability on steep, confined topography. Deep-Sea Research II, this issue [doi:10.1016/j.dsr2.2004.09.011].

Garett, C., 1990. The role of secondary circulation in boundary mixing. Journal of Geophysical Research 95, 3183–3189.

Gerkema, T., Lam, F.-P.A., Maas, L.R.M., 2004. Internal tides in the Bay of Biscay: conversion rates and seasonal effects. Deep-Sea Research II, this issue [doi:10.1016/j.dsr2.2004.09.012].

Gregg, M.C., Kunze, E., 1991. Shear and strain in Santa Monica Basin. Journal of Geophysical Research 96, 16709–16719.

Grimshaw, R.H.J., Khusnutdinova, K.R., 2004. Internal waves in a three-layer bubbly waveguide. Deep-Sea Research II, this issue [doi:10.1016/j.dsr2.2004.09.003].

Hosegood, P., van Haren, H., 2004. Near-bed solibores over the continental slope in the Faeroe-Shetland Channel. Deep-Sea Research II, this issue [doi:10.1016/j.dsr2.2004.09.016].

Isern-Fontanet, J., Font, J., García-Ladona, E., Emelianov, M., Millot, C., Taupier-Letage, I., 2004. Spatial structure of anticyclonic eddies in the Algerian basin (Mediterranean Sea) analyzed using the Okubo–Weiss parameter. Deep-Sea Research II, this issue [doi:10.1016/j.dsr2.2004.09.013].

Munk, W., 1966. Abyssal recipes. Deep-Sea Research 13, 707–730.

Munk, W., Wunsch, C., 1998. Abyssal recipes II: energetics of tidal and wind mixing. Deep-Sea Research I 45, 1977–2010.

St. Laurent, L.C., Nash, J.D., 2004. An examination of the radiative and dissipative properties of the internal tides. Deep-Sea Research II, this issue [doi:10.1016/j.dsr2.2004.09.008].

Simmons, H.L., Hallberg, R.W., Arbic, B.K., 2004. Internal wave generation in a global baroclinic tide model. Deep-Sea Research II, this issue [doi:10.1016/j.dsr2.2004.09.015].

Sutherland, B. R., Flynn, M.R., Dohan, K., 2004. Internal wave excitation from a collapsing mixed region. Deep-Sea Research, this issue [doi:10.1016/j.dsr2.2004.09.002].

van Haren, H., Maas, L., Zimmerman, J.T.F., Ridderinkhof, H., Malschaert, H., 1999. Strong inertial currents and marginal internal wave stability in the central North Sea. Geophysical Research Letters 26, 2993–2996.

Wunsch, C., Ferrari, R., 2004. Vertical mixing, energy and the general circulation of the oceans. Annual Review of Fluid Mechanics 36, 281–314.

ELSEVIER

Deep-Sea Research II 51 (2004) 2889–2904

DEEP-SEA RESEARCH
PART II

www.elsevier.com/locate/dsr2

Internal wave excitation from a collapsing mixed region

B.R. Sutherland*, M.R. Flynn, K. Dohan

Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alta., Canada T6G 2G1

Abstract

The collapse of a uniform density fluid (a "mixed region") into a surrounding ambient fluid with complex stratification is examined by way of laboratory experiments and fully nonlinear numerical simulations. The analysis focuses upon the consequent generation of internal gravity waves and their influence upon the evolution of the collapsing mixed region. In experiments and simulations for which the ambient fluid has uniform density over the vertical extent of the mixed region and is stratified below, we find the mixed region collapses to form an intrusive gravity current and internal waves are excited in the underlying stratified fluid. The amplitude of the waves is weak in the sense that the intrusion is not significantly affected by the waves. However, scaling the results to the surface mixed layer of the ocean we find that the momentum flux associated with the waves can be as large as 1 N/m^2. In simulations for which the ambient fluid is stratified everywhere, including over the vertical extent of the mixed region, we find that internal waves are excited with such large amplitude that the collapsing mixed region is distorted through strong interactions with the waves.
© 2004 Elsevier Ltd. All rights reserved.

1. Introduction

Internal gravity waves (hereafter "IGW") transport energy and momentum vertically through such density stratified fluids as the atmosphere and ocean. IGW generated by flow over topography exert a drag where they break at upper levels in the atmosphere. However, it is only recently that the importance of breaking IGW for transport processes in the near-surface and abyssal ocean has been appreciated.

In the equatorial Pacific, for example, the vertically integrated zonal pressure gradient approximately balances the surface zonal wind stress over seasonal or annual time scales. However, Dillon et al. (1989) have demonstrated that over short time scales the zonal pressure gradient, mesoscale eddy flux and advection cannot be balanced by turbulent stress divergence alone. It was subsequently suggested by Wijesekera and Dillon (1991) and Hebert et al. (1991) that the momentum transported by dynamically generated IGW accounts for this disparity.

*Corresponding author. Tel.: +1 780 492 0573; fax: +1 780 492 6826.
E-mail address: bruce.sutherland@ualberta.ca (B.R. Sutherland).

0967-0645/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dsr2.2004.09.002

Some of the mechanisms for IGW generation below the surface mixed layer which have been investigated through observations, experiments and simulations include dynamic (Kelvin–Helmholtz) instability, turbulent eddies and the "obstacle effect" (e.g. see Moum et al., 1992; Skyllingstad and Denbo, 1994; Sutherland, 1996; Smyth and Moum, 2002; Dohan and Sutherland, 2003). The last of these mechanisms describes wave excitation by the relative horizontal flow of the thermocline underneath the deformed base of the surface mixed layer where the deformation occurs, for example, due to convective plumes or Langmuir circulations. This mechanism is dynamically similar to topographic forcing in which the plumes act as obstacles to a stratified flow.

In the deep ocean, flow over rough bathymetry is perhaps the most important global mechanism for IGW excitation. In particular, observations in the Brazil Basin indicate that IGW generated by barotropic tidal motions over the rough bottom topography of the Mid-Atlantic Ridge are responsible for much of the observed mixing in the region (e.g. Ledwell et al., 2000; St. Laurent et al., 2001). In light of this result, present theories of the global ocean circulation are being revised. It is no longer assumed that the upwelling branch of the thermohaline circulation is a slow diffusive process but rather involves rapid and localized mixing processes that, as yet, are poorly understood.

There are three mechanisms for IGW generation by topographic forcing being actively researched at present: through interaction with bottom topography the barotropic tide may create internal tides (Smith and Young, 2002; Balmforth et al., 2002); quasi-steady flow over topography may generate internal lee waves (e.g. Garrett and St. Laurent, 2002; Sutherland, 2002); and flow near rough topography generates turbulence, which in turn may excite IGW (Sutherland and Linden, 1998; Dohan and Sutherland, 2003).

The last mechanism is particularly complex because IGW can be generated both through interactions with eddies as the turbulent mixed patch develops and through the collapse of the mixing region into the surrounding stratified fluid. This paper aims to elucidate the dynamics governing the collapse of uniform fluid into a stratified ambient and the consequent generation of IGW.

Although this research has been motivated by an oceanographic example, the generic nature of the research allows it to be adapted to a broad range of geophysical circumstances including wave generation due to convective mixing at an atmospheric inversion and wave generation in the stratosphere resulting from high-level thunderstorm outflows.

For example, Fig. 1A schematically represents a circumstance whereby the top of a severe storm spreads horizontally under the tropopause. In addition to the many well-studied mechanisms

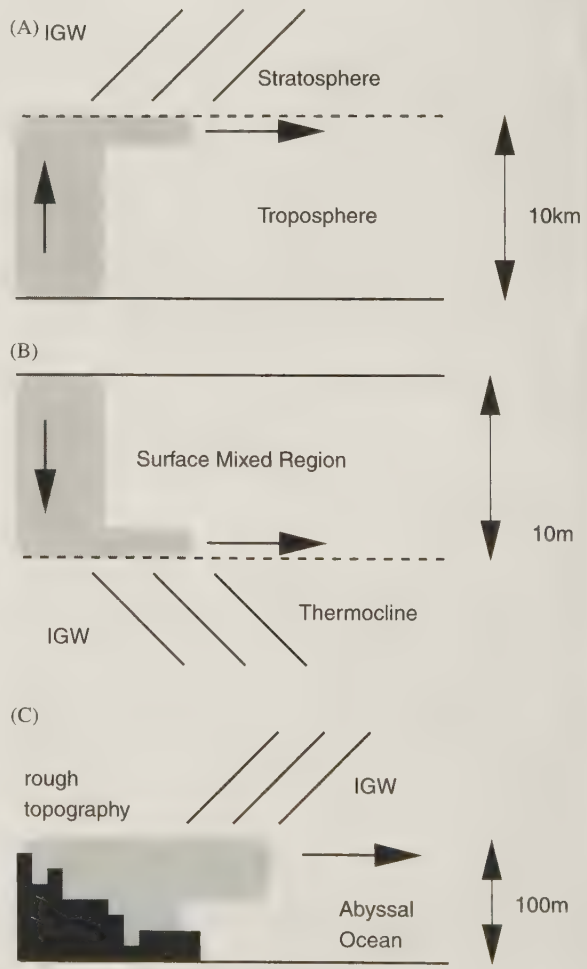


Fig. 1. Schematic showing possible mechanisms for IGW generation in the atmosphere and ocean.

for convective-generation of IGW (Beres et al., 2002; Michaelian et al., 2002) the novel mechanism illustrated here shows IGW excitation in the stratosphere as a result of the horizontal propagation of air moving underneath this stratified region (Flynn and Sutherland, 2004).

Conversely we can imagine an oceanographic circumstance, depicted in Fig. 1B, in which an intrusion moves along the top of the thermocline generating downward propagating IGW. Due to lack of observational data, we can only hypothesize the source of the intrusion; possibly it results from intense localized surface cooling or from storm-induced mixing that extends into the thermocline (Cornillon et al., 1987; Price et al., 1994). It is beyond the scope of this work to examine the origins of the intrusion. Rather we presume that it exists and that it propagates along the thermocline. One goal of this paper is to quantify some aspects of the coupling between this intrusion and the IGW generated in the underlying stratified fluid.

This is an original study in two senses. First, whereas most existing research into intrusions has examined their propagation either in a two-layer or in a uniformly stratified fluid, here we study the hybrid case of an intrusion in a two-layer fluid in which the upper layer has uniform density and the lower layer is uniformly stratified. The first detailed experimental investigation was performed by Flynn and Sutherland (2004), the results of which are adapted to oceanographic circumstances here and are complemented by numerical simulations. Second, whereas most studies of intrusions and gravity currents (which propagate along a horizontal boundary) have analyzed their propagation speeds and mixing efficiencies, here we focus upon the generation of IGW and upon the interaction between these waves and the intrusion.

Geophysical motivation for the study of these interactions is represented by a third scenario shown in Fig. 1C. Here, we imagine that mixing near rough topography in the ocean (whether due to the motion of stratified fluid over a rough surface or due to wave breaking) generates a uniform patch of fluid that collapses to form an intrusion. From the point of view of this study, the important difference between the scenarios in Figs. 1B and C is that the intrusion in the former case propagates primarily within a uniform-density ambient (the surface mixed layer) whereas in the latter case it propagates in a stratified ambient. We will demonstrate that the evolution of the intrusion and the IGW it generates is qualitatively different in these two scenarios.

In Section 2 we describe the setup of laboratory experiments and compare it with the setup of experiments by Maxworthy et al. (2002), who recently studied bottom-propagating currents in uniformly stratified fluid. The experimental results are presented in Section 3. In Section 4 we describe the numerical code used to simulate dynamics close to experimental conditions and their extension to circumstances that cannot be attained in the laboratory. Section 5 directly compares the numerical and experimental results. This section also compares simulations of mixed regions collapsing in a uniformly stratified ambient fluid with those collapsing in a uniform density fluid having an underlying stratified region. In Section 6 we apply our experimental results to a hypothetical situation in which IGW are generated by an intrusion in the ocean's surface mixed layer. Finally, we summarize our results in Section 7.

2. Experimental setup and analyses

Intrusive gravity currents in uniformly stratified fluid have been examined by Wu (1969); Schooley and Hughes (1972); Manins (1976); Amen and Maxworthy (1980); Britter and Simpson (1981); Faust and Plate (1984) and Rooij (1999), and the dynamics of bottom-propagating gravity currents in uniformly stratified fluid have recently been examined by Maxworthy et al. (2002) and Ungarish and Huppert (2002). The gravity currents were created (or simulated) using a lock-release initial condition in which uniform fluid within a lock was released into the stratified ambient.

In all these experiments IGW were generated by two mechanisms: first, through the horizontal motion of the intrusion or gravity current forcing the surrounding stratified fluid and second,

through the flow of ambient fluid replenishing the fluid that had left the lock.

One purpose of our experiments was to reduce the effect of the latter generation mechanism by allowing the lock fluid to collapse into a uniform ambient fluid so that the return flow did not directly excite IGW.

The experimental setup is illustrated schematically in Fig. 2. It shows a two-layer fluid whose upper layer has uniform density and whose bottom layer is uniformly stratified. The lock fluid only spans the depth of the upper layer. Once released, the collapsed lock-fluid propagates along the upper interface of a stratified fluid that moves only in local response to the forward-moving intrusion and not in global response to a return flow.

The second purpose for examining dynamics in such a complex two-layer fluid was that it was representative both of the surface mixed layer of the ocean with an underlying thermocline and, upside-down, of the troposphere underlying the stratosphere. The latter scenario is effectively modelled over vertical distances smaller than the local scale-height of the atmosphere because in the Boussinesq approximation downward-propagating IGW obey the same dynamics as upward-propagating IGW. The stratification prescribed by the former, oceanic, scenario was considered by Delisi and Orlanski (1975); Baines (1982); Thorpe (1998) and Gerkema (2001) in studies of nonlinear interactions between vertically propagating internal waves and interfacial waves. However, their work did not consider the dynamics of intrusions or mean flows.

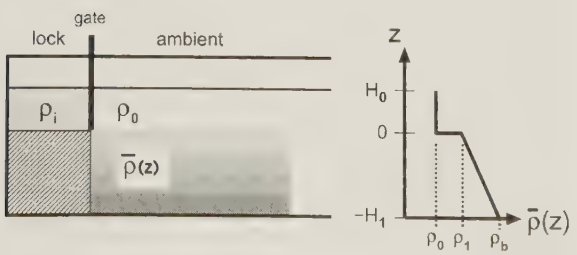


Fig. 2. Schematic of experimental setup showing a lock fluid of density ρ_i behind a gate and overlying a platform (diagonally striped region). The increasing shades of gray to the right of the lock represent the continuous stratification that spans the height of the platform.

Our experiments were performed in a glass tank measuring 197.1 cm long by 19.9 cm wide by 48.5 cm tall. A 30.0 cm tall and 19.6 cm long platform that spanned the width of the tank was installed at one end.

The tank was filled with a two-layer fluid in which the lower layer was stratified and the upper layer had uniform density, as illustrated schematically in Fig. 2. This was established in a two-step process. First, a uniformly stratified fluid was added to the tank using the standard "double bucket" technique (Oster, 1965). The top of this stratified layer was at 30 cm, equal to the level of the top of the platform. Using a sponge-float, a 15.0 cm deep layer of fresh or slightly saline water of density ρ_0 was then layered on top of the uniformly stratified layer.

The resulting density profile was measured using a vertically traversing conductivity probe (Fast Conductivity and Temperature Probe—Precision Measurement Engineering). From the vertical density profile in the lower layer, $\bar{\rho}(z)$, we compute the squared buoyancy frequency $N^2 = -(g/\rho_{00})\,\mathrm{d}\bar{\rho}/\mathrm{d}z$, in which g is the acceleration due to gravity and ρ_{00} is a characteristic density of this Boussinesq fluid. The conductivity probe also measured the density change, $\rho_1 - \rho_0$ and thickness, H_{Int}, of the step between the stratified and uniform fluids. Here ρ_1 is the density of the fluid at the top of the stratified layer. Typically we found $H_{\mathrm{Int}} \simeq 1$ cm.

After filling the tank a 0.4 cm thick gate was inserted between a pair of vertical glass guides and was lowered to the level of the platform near its edge, thus forming a 15 cm deep lock of length $\ell = 18.6$ cm. Salt and a small volume of food-coloring were then added to the lock fluid after which it was mixed vigorously for approximately 15 s. The resulting homogeneous lock fluid had density ρ_i.

In total, 39 experimental runs were performed in which the following three parameters were varied: N^2, $\sigma_{10} = (\rho_1 - \rho_0)/\rho_0$ and $\sigma_{i0} = (\rho_i - \rho_0)/\rho_0$. Experiments were performed with N^2 ranging from 0.37 to 1.21 s^{-2}, σ_{10} ranging from 0.0015 to 0.0102 and σ_{i0} ranging from 0.0005 to 0.0023. In all but

three cases, we set $\sigma_{i0} < \sigma_{10}$ so that the intrusion density matched an isopycnal surface within the interface of the two-layer fluid. Thus we were able to examine the response of the stratified layer beneath the intrusion which, in itself, did not penetrate deeply into the lower layer.

Runs began by rapidly removing the gate from the tank. The lock fluid then collapsed to form an intrusion that propagated along (or in the three cases for which $\sigma_{i0} > \sigma_{10}$, just below) the interface. A CCD camera was used to record the resulting fluid motions. It was placed $L_c = 400\,\text{cm}$ in front of the tank and captured a field of view that covered a horizontal distance $L_w = 87\,\text{cm} \simeq 4.7\ell$. In all cases, the edge of the platform was at the far left extremity of the field of view. Images were analyzed using DigImage (Dalziel, 1992). It was used to convert "snapshot" images of the experiment (having a spatial resolution of approximately $0.2\,\text{cm} \times 0.2\,\text{cm}$) into time series images (having a temporal resolution as small as 1/30 s).

The IGW that appeared in the (stratified) lower layer were visualized using "synthetic schlieren" (Sutherland et al., 1999; Dalziel et al., 2000), a technique that exploits variations in the index of refraction with salinity in order to visualize and measure the deflection of light rays as they travel through stratified fluid of spatially and temporally varying density gradient. By measuring the deflection one can determine the structure and amplitude of the waves continuously and non-intrusively.

Specifically, we measure the value of $\Delta N^2(x, z, t)$, the perturbation to the squared buoyancy frequency, and its time derivative N_t^2. The amplitudes of these fields are themselves proportional to the maximum vertical displacement of the waves, A.

IGW were observed in all but three experiments in which the stratification, $N^2 = 0.37\,\text{s}^{-2}$, was weakest. For these, the absence of observed waves was attributed to the signal from the waves being weaker than the signal noise, for example, due to temperature fluctuations in the laboratory and degradation of the transmitted images. Signal noise is less significant in experiments with stronger stratification because, for a wave of fixed amplitude and spatial structure, the wave signal increases as N^3.

3. Experimental results

3.1. Intrusion analyses

Once released from the lock the gravity current propagates at approximately constant speed, U_i, over the field of view of the camera. This speed was measured from horizontal time series of the digitized experimental images (e.g. see Mehta et al., 2002).

The intrusion speeds measured in all experiments are plotted in Fig. 3A. Here U_i is normalized by a characteristic gravity current speed $C_0 = \sqrt{g\sigma_{i0}H_0}/2$. This is the predicted speed the intrusion would have if it was energy-conserving and if the level $z = 0$ was a rigid bottom (Benjamin, 1968). In the limit $\sigma_{i0} \ll 1$, the intrusion indeed propagates primarily within the uniform density fluid and the underlying stratified fluid acts approximately as a rigid bottom. We expect, and show, that increasing discrepancies in the predicted intrusion speed occur as σ_{i0} increases.

The plot shows that in most experiments the normalized intrusion speed is moderately smaller than 1.0 and that the relative speed generally decreases as σ_{i0} increases. Conversely, Fig. 3B shows that at fixed σ_{i0} (in this case $\simeq 0.0005$) the relative intrusion speed increases as σ_{10} increases. This is true for other values of σ_{i0} we examined. In cases with small σ_{i0}, the speed lies within 10% of the value C_0 and the specific value of N^2 has a relatively small effect on U_i.

We conclude that, at least in the parameter ranges we have examined, the intrusion speed is governed primarily by energy-conserving gravity current dynamics and that deviations from C_0 occur because the intrusion is able to slump below the level $z = 0$ and so effectively reduce its height in the ambient above $z = 0$. Larger σ_{i0} leads to deeper slumping and larger σ_{10} inhibits slumping. These observations are consistent with measurements of the current depths above and below the interface (Flynn and Sutherland, 2004).

The horizontal extent of the intrusion head, l_i, was determined from vertical time series taken

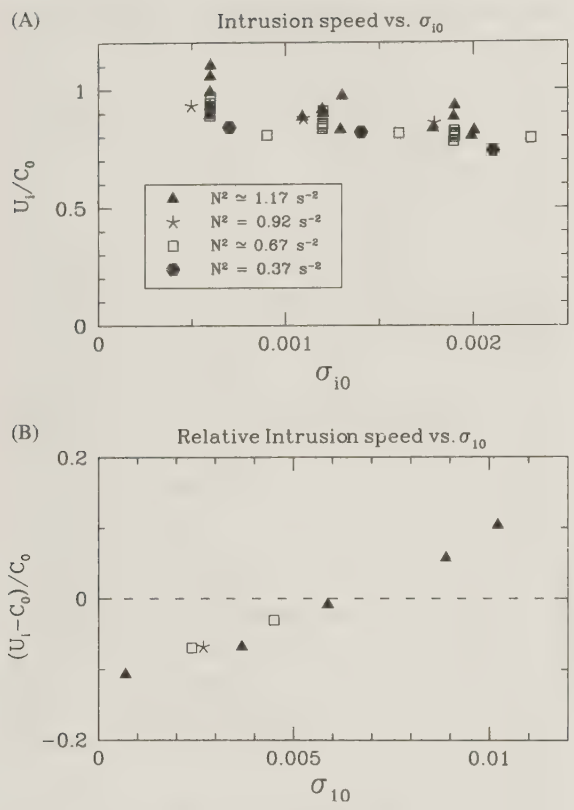


Fig. 3. (A) Normalized gravity current speeds, U_i, as a function of σ_{i0} for all experiments and (B) $U_i - C_0$ as a function of σ_{10} for experiments with small $\sigma_{i0} \simeq 0.0005$. In both plots, the speeds are normalized by $C_0 = (g\sigma_{i0}H_0)^{1/2}/2$. Different symbols are plotted depending on the value of N^2 below $z = 0$, as indicated in the legend.

one lock length from the gate. Specifically, l_i corresponds to the distance between the leading edge of the intrusion and the projected location at which the height of the intrusion head is equal to the height of the trailing tail. It was difficult to measure l_i precisely because the rear of the intrusion head was often obscured by transient billows. Typical measurement errors were 15%.

Knowing U_i and l_i, we estimate the characteristic frequency, ω_i, associated with the forcing on the stratified fluid imparted by the intrusion head by

$$\omega_i = U_i \frac{2\pi}{l_i}. \tag{1}$$

3.2. Internal gravity wave analyses

When the intrusion is released it excites IGW in the stratified fluid beneath it. Synthetic schlieren measures the corresponding ΔN^2 and N_t^2 fields, which are shown for a particular experiment at time $t = 15$ s in Fig. 4. Only the snapshot in the camera's field of view between $x = 25$ and 90 cm is shown. The position $x = 0$ corresponds to the end wall of the tank so that the left extremity of the image is 6.8 cm from the original position of the gate. The vertical range of the images extends from $z = -20$ cm (10 cm above the bottom of the tank) to $z = 0$ cm, the level of the interface between the stratified and unstratified fluid.

As evident from images above $z = 0$ (not shown) the front of the intrusion is situated near $x = 60$ cm at this time. The ΔN^2 field, which is proportional to the change in density gradient, shows a disturbance propagating rapidly upstream just below the level of the interface. IGW propagating downward directly below the intrusion are also evident. ΔN^2 is large near $z = 0$ because the background N^2 is relatively large at the interface. Below the interface ΔN^2 is as large as $0.3\,s^{-2}$. That is, the background N^2 changes by as much as 25% due to the stretching and compressing of isopycnal surfaces by waves.

Because it takes a time derivative, the N_t^2 field filters slowly varying disturbances and tends to enhance visualisation of vertically propagating IGW rather than long horizontal wavelength (columnar) disturbances. This is evident in Fig. 4B, where downward left-to-right-propagating waves are shown.

The properties of downward-propagating IGW were estimated from equally spaced vertical time series of N_t^2, from which horizontal time series were then constructed. Power spectra were then used to determine the wavenumber, (k_x, k_z), and frequency, ω, of the waves containing most of the energy. In practise we find this peak is quite sharp indicating that the waves are generated within a narrow frequency and wavenumber band (Flynn and Sutherland, 2004).

Fig. 5A plots the measured peak values of frequency and horizontal wavenumber. We do not

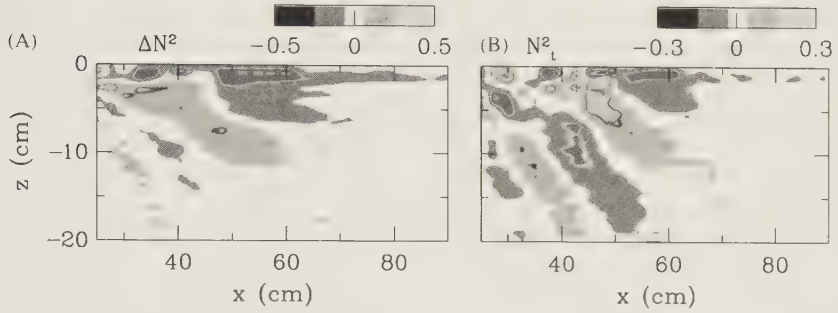


Fig. 4. (A) ΔN^2 and (B) N_t^2 fields at $t = 15$ s in an experiment with $N^2 = 1.17\,\mathrm{s}^{-2}$ below $z = 0$, $\sigma_{10} = 0.0037$ and $\sigma_{i0} = 0.0020$.

plot the vertical wavenumber because the wavelength was found to be smaller than but comparable to the vertical extent of the field of view. This introduced unacceptably large errors in the estimates of k_z. The horizontal wavenumber is plotted normalized by $k_i = 2\pi/l_i$, the estimated wavenumber of the forcing imparted by the intrusion head of length l_i. Likewise the frequency is plotted normalized by $\omega_i = U_i k_i$, the estimated forcing frequency.

The plot shows that k_x is moderately larger than k_i. Although there is quite a bit of scatter in the data, much of this can be attributed to the 15% error associated with the estimate of l_i. We conclude that the extent of the intrusion head is primarily responsible for setting the horizontal scale of the waves.

The plot also shows that ω is consistently larger than ω_i. Indeed, the wave frequency is up to 3 times larger than the equivalent forcing frequency. The scatter in the points in this case well exceeds the associated error bars, and so we conclude that the forcing frequency is not primarily responsible for setting the wave frequency.

Rather it appears that the buoyancy frequency, N, determines the value of ω. This is demonstrated in Fig. 5B. Here we plot the relative amplitude A of the waves against ω/N. We find that the frequency ratio generally lies between values of 0.5 and 0.75 for a wide range of forcing frequencies and values of N. From the linear dispersion relation of IGW, we know $\Theta = \cos^{-1}(\omega/N)$ represents the angle formed by lines of constant phase with the vertical. Thus we find that the waves generally propagate at angles between 41° and 60°.

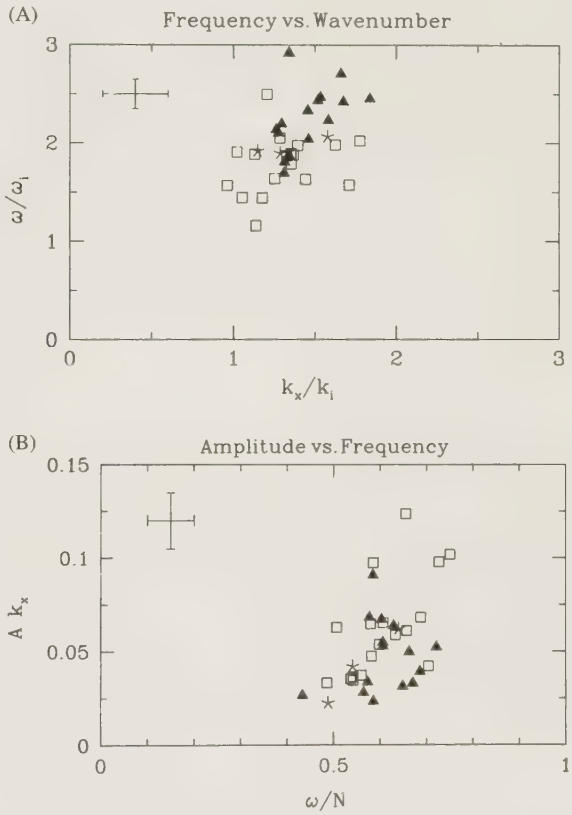


Fig. 5. (A) Measured frequency and wavenumber of IGW expressed as fractions of forcing frequency, $\omega_i = U_i(2\pi/l_i)$, and wavenumber, $k_i = 2\pi/l_i$, respectively. (B) Relative amplitude plotted as a function of wave frequency relative to the background buoyancy frequency. Typical error bars are indicated in the top left-hand corner of both plots. Symbols are plotted as indicated in Fig. 3.

This observation is consistent with Dohan and Sutherland (2003) who found that IGW generated below a turbulent, but stationary, mixed region

propagated within a narrow frequency band. They proposed that when IGW are generated by a fluidic, rather than rigid disturbance, they resonantly interact with the source region in a manner that most strongly excites those waves that vertically transport the most horizontal momentum. These waves, which propagate at angles of about 45°, exert the greatest influence upon the mixing region by exerting the greatest drag to it as a consequence of momentum extraction.

The wave amplitude, A (the maximum vertical displacement), is determined from $A_{N_t^2}$, which in turn is found by computing half the maximum difference between adjacent troughs and crests in vertical time series images of the N_t^2 field. To avoid initial transient effects, the time series were taken one lock-length from the gate. Measurements were typically made 3 cm below the interface. This level was chosen as the approximate position close to the source region but not so close that quantitative schlieren measurements were unreliable due to large isopycnal displacements creating overly distorted images.

4. Numerical simulation setup

The collapse and propagation of the gravity current and the dynamics of the IGW it generates was simulated numerically using a code that solved the fully nonlinear Navier–Stokes and mass-conservation equations in two dimensions. The equations were simplified by invoking the Boussinesq approximation, which assumes that the maximum change of salt-water density from the top to bottom of the tank is negligibly small compared with the density of water itself.

Thus the code solved the following two coupled partial differential equations for the basic-state fields of vorticity, ζ, and perturbation density, ρ':

$$\frac{\partial \zeta}{\partial t} = -u\frac{\partial \zeta}{\partial x} - w\frac{\partial \zeta}{\partial z} + \frac{g}{\rho_{00}}\frac{\partial \rho'}{\partial x} + \nu\nabla^2\zeta \tag{2}$$

and

$$\frac{\partial \rho'}{\partial t} = -u\frac{\partial \rho'}{\partial x} - w\frac{\partial \rho'}{\partial z} - w\frac{\mathrm{d}\,\bar{\rho}}{\mathrm{d}\,z} + \kappa\nabla^2\rho'. \tag{3}$$

Here the horizontal and vertical components of the velocity field were given by u and w respectively. These were related to the streamfunction, ψ, by $u = -\partial\psi/\partial z$ and $w = \partial\psi/\partial x$, and ψ, in turn was related implicitly to the vorticity field by

$$\zeta = -\nabla^2\psi. \tag{4}$$

Thus, once ζ was determined at a particular time, the Poisson equation, 4, was solved to find ψ from which (u, w) was found and ultimately used in (2) and (3) to determine how the basic-state fields changed in time.

The physical parameters of the system of equations were g, the gravitational acceleration, ρ_{00}, the characteristic density, ν, the kinematic viscosity, and κ, the diffusivity of salt water. Assuming weakly saline solutions, we took each of these to be constant. Specifically, $g = 980\,\mathrm{cm/s^2}$, $\rho_{00} = 1\,\mathrm{g/cm^3}$, $\nu = 0.01\,\mathrm{cm^2/s}$ and $\kappa = 0.01\,\mathrm{cm^2/s}$. The last of these was unphysical; $10^{-5}\,\mathrm{cm^2/s}$ would have been more appropriate for the diffusivity of salt water. If we used this value, however, the simulation would have been numerically unstable. For the purposes of simulating the experiments, the diffusivity term in 3 acted solely as a filter to damp small-scale numerical noise. The value we used was still sufficiently small that the essential dynamics of the gravity current's propagation and wave generation were not affected.

The code was originally designed to model the evolution of unstable stratified parallel flows (Smyth and Peltier, 1991), and so it solved the equations in a L-long by H-deep domain which was a horizontally periodic channel with free-slip upper and lower boundary conditions.

To adapt the code to model the experiment described in Section 2, we chose a computational domain with twice the horizontal extent of the tank (i.e. $L = 2L_T$) and we prescribed initial conditions with reflection symmetry about the mid-vertical plane of the domain. We confirmed, a posteriori that indeed there was no significant motion across this reflection plane nor across the right and left-extremes of the domain. Hence, the right-half of the domain simulated the dynamics of a gravity current bounded on four sides by free-slip boundaries. Simulations were run with $H = H_T = H_0 + H_1$, the depth of the tank, and with

$H = 4H_T$. The latter enabled us to simulate the experiment while eliminated the effect of bottom reflection of the IGW over the duration of interest.

Initial conditions for a particular simulation are shown in Fig. 6. The right-hand plots show the initial profiles of the background density, $\bar{\rho}(z)$ and the corresponding squared buoyancy frequency $N^2(z)$. Explicitly,

$$N^2(z) = \begin{cases} N_0^2, & 0 \leqslant z \leqslant H_0, \\ N_{\rm Int}^2, & -H_{\rm Int} \leqslant z < 0, \\ N_1^2, & -H_1 \leqslant z < -H_{\rm Int}. \end{cases} \quad (5)$$

In particular, as illustrated in Fig. 6, when $N_0^2 = 0$ the background density was constant ($= \rho_0$) over the top H_0 of the domain. Below a $H_{\rm Int} = 1$ cm deep interface the density increased linearly as characterized by a constant squared buoyancy frequency equal to N_1^2. Across the interface the relative density increased by σ_{10}, which was written in terms of $N_{\rm Int}^2$ by $\sigma_{10} = H_{\rm Int} N_{\rm Int}^2 / g$.

As well as the profiles shown, we ran simulations with $N^2(z) = N_1^2$ everywhere and with $N^2(z) = N_1^2$ above and below a $H_{\rm Int} = 1$ cm deep interface with the density increasing linearly with depth by $\rho_1 - \rho_0 = 0.0037\,{\rm g/cm^3}$ across the interface.

These profiles determined the density structure at every horizontal position along the domain. In addition, we superimposed an initial density perturbation such that

$$\rho'(x, z, t = 0) = \begin{cases} \rho_i, & |x| < \ell, 0 \leqslant z \leqslant H_0, \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

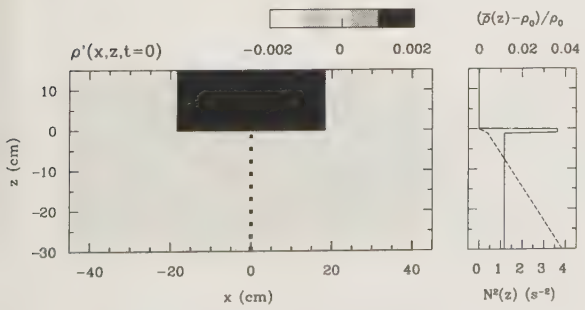


Fig. 6. Perturbation density field (left), ambient density profile (right dotted) and N^2 profiles (right, solid) used to initialize the numerical simulations.

The density perturbation field is shown by color contours on the left-hand side of Fig. 6 for the case $\rho_i = 0.0020\,{\rm g/cm^3}$ and $\ell = 18.6$ cm. Here, only a window within the actual domain size is shown.

Because of the additional complexity involved with the numerics, the simulation did not model the effect of a rigid boundary beneath the lock fluid. Rather, stratified ambient fluid underlies the lock and is free to move in response to the collapse of fluid in the lock. The simulations thus enable us to examine the effect of this response to wave generation.

The details of the numerical representation of the basic-state equations are given by Sutherland and Peltier (1994). Here we ran simulations with a typical horizontal and vertical resolution of 0.2 and 0.06 cm, respectively. The time step was 0.005 s.

5. Simulation results

5.1. Direct simulation of a laboratory experiment

We have run simulations using parameters nearly identical to that of the particular laboratory experiment, shapshots of which are shown in Fig. 4. Explicitly, we initialize the simulation with an ambient N^2 profile given by 5 with $N_0^2 = 0$, $N_{\rm Int}^2 = 3.63\,{\rm s^{-2}}$, $N_1^2 = 1.17\,{\rm s^{-2}}$ and with $H_0 = 15$ cm, $H_{\rm Int} = 1$ cm and $H_T = 45$ cm. The buoyancy frequency at the interface ($-1 \leqslant z < 0$) corresponds to a relative density increase of $\sigma_{10} = 0.0037$. An initial density perturbation is given by (6) with $\rho_i = 0.0020\,{\rm g/cm^3}$. This is applied in a rectangular region bounded by $0 \leqslant z \leqslant H_0$ and $|x| \leqslant \ell$.

Snapshots taken from this simulation of the perturbation density and vorticity fields are shown in Fig. 7. Though the total horizontal extent of the computational domain is $|x| < L_T$ with $L_T = 197.1$ cm, only the range $0 \leqslant x \leqslant 90$ is shown. The images clearly show that the lock fluid collapses to form an intrusive gravity current that propagates along the interface between the upper uniform density ambient and the lower stratified ambient fluid. Consistent with experiments, Kelvin–Helmholtz billows develop behind the head of

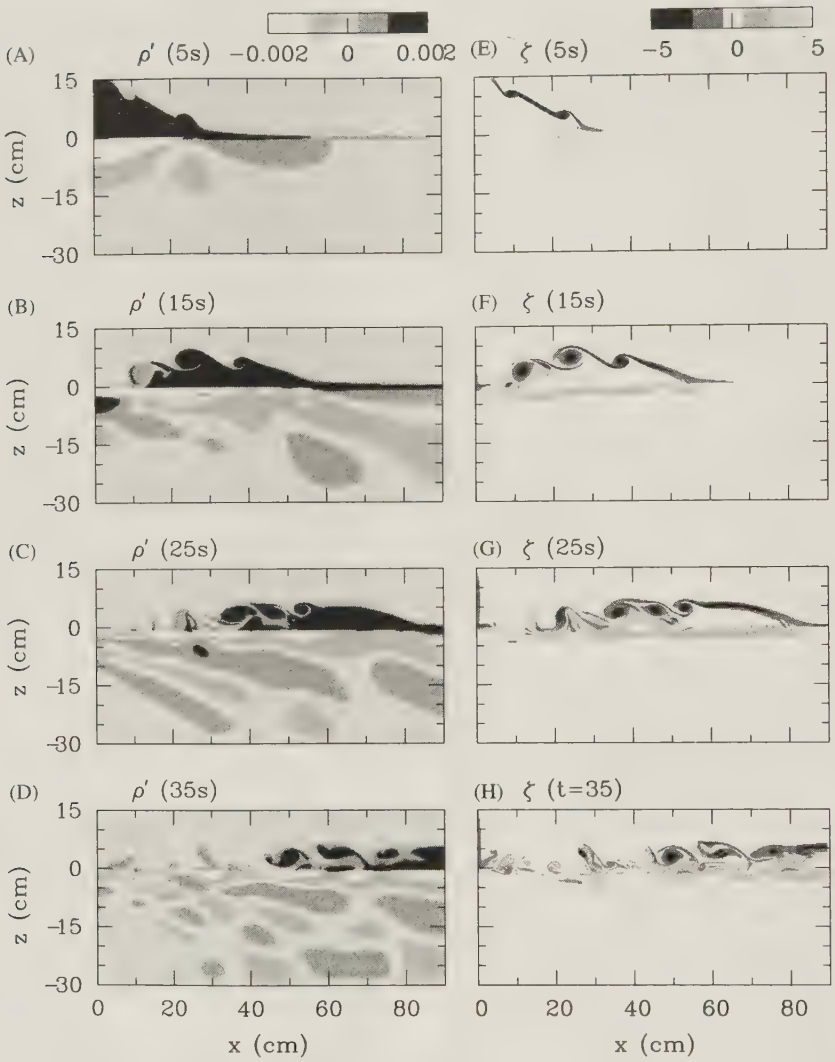


Fig. 7. Snapshots from numerical simulation with $N_0^2 = 0$, $N_1^2 = 1.17\,\mathrm{s}^{-2}$, $\sigma_{10} = 0.0037$ and $\sigma_{i0} = 0.0020$. The perturbation density field ρ', like σ_{i0}, is given in nondimensional units. The vorticity field ζ' is given in units of s^{-1}.

the intrusion on its upper flank, but such instability is not apparent on its lower flank. The mixing occurs primarily behind the head of the intrusion so that the density near the nose of the intrusion remains constant. Likewise, the vorticity field remains zero near the nose except on thin sheets surrounding the gravity current head. Thus, consistent with the experimental observations of Lowe et al. (2002), the velocity near the nose of the current is approximately constant.

IGW are excited immediately after the fluid begins to collapse. The waves transport energy down and rightward until they reflect off the bottom at $z = -30$. This reflection is evident at $t = 25\,\mathrm{s}$ after which time the wave field exhibits a complicated pattern of interference between downward and upward propagating waves.

We have performed a simulation (not shown) in which all parameters are identical to those in the above simulation except that we quadruple the

depth of the domain while keeping the vertical resolution the same. This simulation gives the same results at early times, but reflection does not corrupt the wavefield at $t = 25$ s.

Corresponding to Figs. 7B and F, Fig. 8 plots the ΔN^2 and N_t^2 fields at $t = 15$ s. For comparison with the experimental results shown in Fig. 4, the fields are shown only for $25 \leqslant x \leqslant 90$ and $-20 \leqslant z \leqslant 0$.

Both the simulations and experiments show left-to-right downward-propagating IGW emanating from the head of the intrusion whose front is near $x = 60$ cm. However, the amplitude of the simulated ΔN^2 field is almost twice that measured in experiments. Furthermore, whereas in the experiments a regular pattern of waves is evident in the N_t^2 field (Fig. 4B), no such regularity is evident in the simulated field. Rather the lines of constant phase vary from nearly vertical to nearly horizontal and the waves themselves appear to emanate from a localized source near $(x, z) \simeq (55\,\text{cm}, -2\,\text{cm})$.

We attribute these discrepancies to the fact that the simulations do not include the effect of a rigid bottom boundary directly beneath the lock as would be present due to the platform employed in the experiments. Fig. 7A shows that, without the platform, fluid beneath the lock is displaced downward during the initial collapse phase. The associated values of ΔN^2 at this location (not shown) are an order of magnitude larger than those associated with waves in the experiments. This could explain in part the occurrence in the simulations of large-amplitude disturbances of long horizontal extent at $t = 15$ s (Fig. 8A). Supporting this assertion, we find that the N_t^2 field (which filters long-time scale shallow motions) has comparable simulated and experimental amplitude for the primary wave (shown by light-gray contours) below the intrusion front.

The structure of the waves in the two cases is nonetheless different. We attribute the discrepancy to the approximately horizontal flow induced by the initial displacement of fluid below the lock in simulations. This flow effectively Doppler shifts the transmitted waves and prevents them from being transmitted far away from the source. Indeed this Doppler shifting is apparent in the waves shown in Fig. 8B. The primary wave and trailing trough have nearly vertical phase lines for $-10 \leqslant z \leqslant -5$ but these become more shallow below $z \simeq -10$ cm. Further behind the intrusion front, the waves generated by the source are sufficiently Doppler-shifted that the waves are evanescent. Diagnosis of the horizontal velocity field at $t = 15$ s (not shown) reveals that the flow speed decreases by approximately 2 cm/s from $z = 0$ to -5 cm. For waves with horizontal wavenumber $2\pi/20\,\text{cm}^{-1}$ and intrinsic frequency $0.6N$ (typical values taken from experimental data), the Doppler-shifted frequency at $z = -5$ cm would be $1.2N$, which demonstrates that the waves are indeed evanescent.

These simulations thus serve to demonstrate that the mechanism for wave generation sensitively depends upon the flow in the ambient. This observation is consistent with the conclusions by Hughes (1996) who examined in experiments the collapse of a mixed region in a uniformly stratified

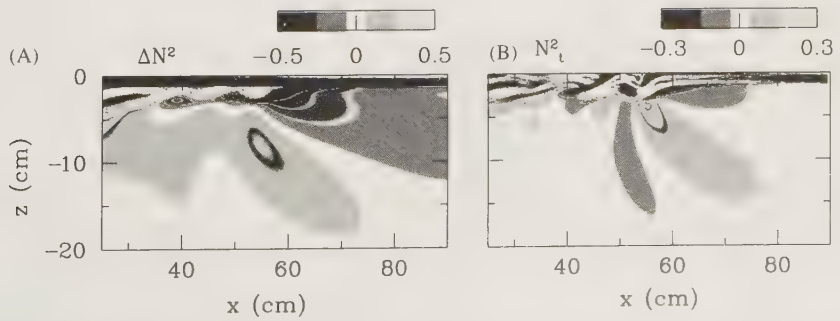


Fig. 8. Snapshots of (A) ΔN^2 and (B) N_t^2 fields at $t = 15$ s in a simulation with the same parameters as those given in Fig. 7.

fluid. He found that vertically propagating IGW were prevented from propagating away from the intrusions on either side of the mixing region because the return flow above and below the intrusion had sufficiently large velocity to Doppler-shift waves to evanescent frequencies.

5.2. Simulation of uniformly stratified ambient

One reason for the relatively complicated experimental setup we have used is that we wanted the lock-fluid to collapse within an ambient fluid of uniform density over the depth of the lock so that the process of lock-collapse and, in particular, the development of a return flow into the lock would not in itself directly excite IGW. IGW can only be excited as a consequence of the intrusion moving along the interface above a stationary stratified ambient.

In order to demonstrate the effect of a lock-fluid collapsing in a stratified ambient we present here the results of two simulations which are simple extensions of the simulation described in Section 5.1.

In both, we set the ambient fluid to be uniformly stratified with $N^2 = 1.17\,\mathrm{s}^{-2}$ both above $z = 0$ and below $z = -H_{\mathrm{Int}}$. For $-H_{\mathrm{Int}} \leqslant z < 0$, the stratification is the same as before. A density perturbation is introduced in a model lock above $z = 0$ and for $|x| < \ell$.

In one simulation, the lock has the same stratification as the ambient, but a density perturbation of $\rho_i = 0.0020\,\mathrm{g/cm}^3$ is superimposed. The results of this simulation are shown in Fig. 9, and should be compared with the top four images in Fig. 7. The perturbation density field at times 5 and 15 s (Figs. 9A and B, respectively) demonstrates notably different behavior from that in the simulation with a uniform ambient. Here the density anomaly does not remain concentrated in a collapsing current. Rather, it creates a disturbance that propagates rapidly to the right at the same vertical level as the lock. After 15 s the perturbation density field is characterized by IGW throughout the domain. No sharp front, as one would associate with a gravity current head, is evident. Likewise the vorticity field (Figs. 9C and D, respectively) exhibits no characteristic traits of Kelvin–Helmholtz instability. Indeed, the vorticity field is significantly weaker as evident from the range of grayscale which is five times smaller than that used to plot Figs. 7E and F.

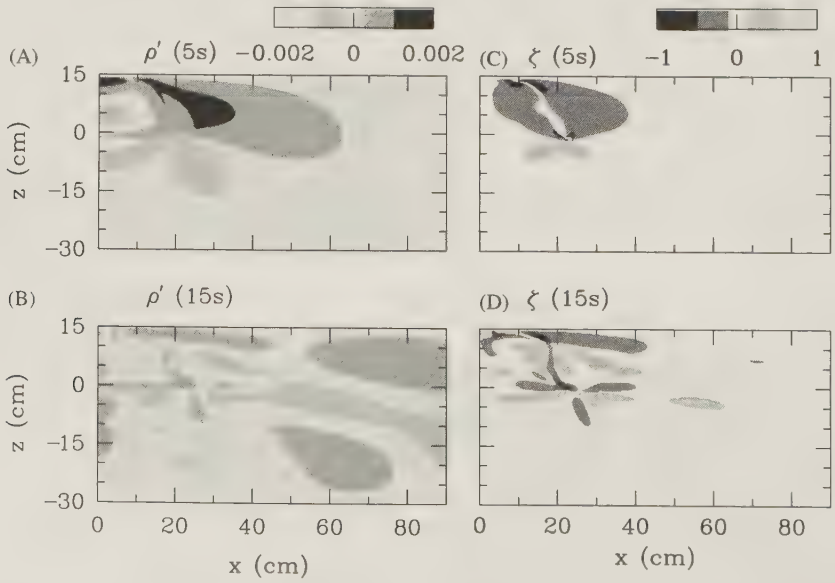


Fig. 9. Snapshots from a numerical simulation with an almost uniformly stratified ambient (see text) and with a lock-fluid having the same density gradient as the ambient but with a density difference $0.0020\,\mathrm{g/cm}^3$ at each vertical level in the lock.

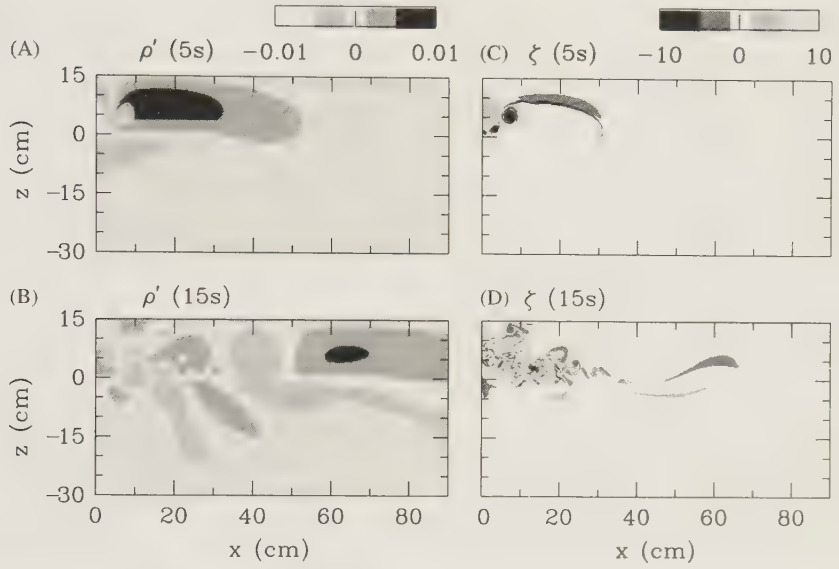


Fig. 10. Snapshots from a numerical simulation with an almost uniformly stratified ambient (see text) and with a uniform initial lock-fluid density.

In a second simulation, the lock has uniform density which is greater than the density of the ambient at $z = 0$ by $0.0020\,\mathrm{g/cm^3}$. The results of this simulation is shown in Fig. 10. Again a collapsing gravity current head is not evident. Rather, IGW are found to propagate rapidly to the right. The vorticity field is stronger in this case due to baroclinically generated negative vorticity in the lock during the initial stages of collapse (Fig. 10c). However, the initially coherent vortices are rapidly broken down into smaller-scale structures in the lock after 15 s.

Both simulations show that the flow dynamics change qualitatively when the ambient fluid throughout the depth of the lock is stratified rather than unstratified.

6. Application to surface mixed region

Here we adapt the experimental results in order to form order-of-magnitude estimates of the momentum flux associated with waves generated by intrusions propagating in the surface mixed region above the thermocline. The approach is analogous to that taken by Flynn and Sutherland (2004) in their study of IGW generated in the stratosphere by thunderstorm outflows below the tropopause.

The ability of IGW to vertically transport horizontal momentum is characterized by the average Reynolds stress, $\rho_{00}\langle uw\rangle$. This in turn can be related to experimental parameters in an empirically determined formula (Flynn and Sutherland, 2004):

$$\frac{\rho_{00}\langle uw\rangle}{(N_1 l_i)^2} = 10^{-3.7\pm0.3}\rho_{00}\left(\frac{U_i}{\sqrt{\sigma_{10} g h_i}}\right)^{2.9\pm0.6}. \qquad (7)$$

The exponents with associated errors are determined from the slope and intercept of the best-fit line through log–log plots of momentum flux and current speed measurements gathered from 30 experiments. Here h_i is the total height of the intrusion head, which observations show can be related to the length of the head by $l_i \simeq 4.7 h_i$.

Because there are no direct observations of intrusions above the thermocline, we must estimate values of the parameters in (7) in order to form an estimate for the momentum flux. We do so by focusing on a particular geophysical event: the passage of Hurricane Norbert over the eastern north Pacific Ocean during September and October of 1984. The observations and simulations

of Price et al. (1994) give estimates of the surface cooling and mixed layer deepening due to this storm. As a result of storm-induced vertical mixing, they predict a deepening of the surface mixed layer from 50 m to approximately 90 m and a corresponding mixed layer density increase of approximately $0.64\,\mathrm{kg/m^3}$.

Relative to the unperturbed ambient, we speculate that this mixed parcel of fluid ultimately collapses and propagates as a fluid intrusion at a depth $z_{\mathrm{NB}} = -58.4\,\mathrm{m}$. The intrusion's plane of propagation is therefore directly below the level $z = -50\,\mathrm{m}$, which denotes the bottom of the uniform ambient mixed region.

To estimate the relevant parameters associated with this intrusion, we define σ_{10} such that

$$\sigma_{10} = \frac{Ch_i}{\rho_{00}}\left(-\frac{\mathrm{d}\bar{\rho}}{\mathrm{d}z}\right), \tag{8}$$

where $-\mathrm{d}\bar{\rho}/\mathrm{d}z$ is the ambient density gradient between $z = -50\,\mathrm{m}$ and $-90\,\mathrm{m}$ and

$$C = \min\left[\frac{(z_{\mathrm{NB}} - z_0) + h_1}{h_i}, 1\right]. \tag{9}$$

Here $z_0 = -50\,\mathrm{m}$ and h_1 is the intrusion's depth of penetration into the (lower) stratified layer. The constant, C, is defined so that σ_{10} represents the magnitude of the normalized density jump across a distance Ch_i.

We further relate h_i and σ_{10} using the first-order perturbation approximations for the intrusion height in a two-layer model (Holyer and Huppert, 1980; Flynn and Sutherland, 2004). Thus we find

$$\frac{h_i}{z_{\mathrm{NB}}} = \frac{1}{2} + \frac{5}{32}\left(\frac{\sigma_{i0}}{\sigma_{10}}\right). \tag{10}$$

In the wide range of experiments by Flynn and Sutherland (2004), this estimate was found to be moderately smaller than the observed heights, but was sufficiently close that its use here is appropriate to determine order-of-magnitude momentum flux estimates. Eqs. (8)–(10) are solved using a simple iterative routine from which we find $\sigma_{10} \simeq 0.0010$. A posteriori this value of σ_{10} is sufficiently small that such an order-of-magnitude estimate is self-consistent.

Using this value of σ_{10}, we predict the vertical and horizontal extent of the intrusion head to be $h_i \simeq 35.1\,\mathrm{m}$ and $l_i \simeq 166\,\mathrm{m}$, respectively. We estimate the propagation speed from the first-order perturbation expansion of the predicted steady-state intrusion speed in a two-layer fluid (see Eq. (2.8) of Flynn and Sutherland, 2004). The normalized speed of propagation is thereby determined to be $U_i/\sqrt{\sigma_{10} g h_i} \simeq 0.5$, and so we find $\rho_{00}\langle uw\rangle$ ranges between 0.2 and $2\,\mathrm{N/m^2}$.

By comparison, IGW excited below the mixed region of the equatorial Pacific were observed to have typical Reynolds stress values of approximately $0.3\,\mathrm{N/m^2}$ (Lien et al., 1996), and Wijesekera and Dillon (1991) estimated that for narrow-band, dynamically generated IGW, the associated Reynolds stress ranged from 0.1 to $0.6\,\mathrm{N/m^2}$.

The above results suggest that the local influence of IGW excited by fluid intrusions is roughly comparable to those generated by other near-surface mechanisms. Globally, however, their significance is likely to be small because severe storms of sufficient energy to substantially deepen the surface mixed layer are relatively intermittent.

7. Conclusions

We have shown that intrusions which propagate over the thermocline can act locally as a significant source of IGW. However, the momentum flux associated with the waves and, indeed, whether the waves are generated at all depends sensitively upon the flow speed of the stratified fluid. Whether from the initial transient collapse of a mixed region into the stratified fluid below or from the horizontal flow of stratified fluid into a collapsing mixed region, long horizontal scale disturbances can be excited. These propagate rapidly away from the source and can filter vertically propagating waves.

We have adapted our experimental results to the ocean surface, but it is evident that IGW generation from a collapsing mixed region would likewise occur in the abyssal ocean. Such mixing could result directly from turbulence generated by flow over rough topography or indirectly from the breaking of inertia gravity waves. In either case, our simulations show that a mixed region in a stratified ambient does not collapse in the form of a standard interfacial gravity current. Rather it

excites IGW of such large amplitude that the mixed region is itself significantly distorted by the waves. The details of these dynamics including the long-term effects of rotation are presently under investigation.

Acknowledgements

This work was financially supported by NSERC's PGS A and Discovery Grants programs, the Killam Trust, Alberta's iCore Program and by the Canadian Foundation for Climate and Atmospheric Sciences (CFCAS).

References

Amen, R., Maxworthy, T., 1980. The gravitational collapse of a mixed region into a linearly stratified solution. Journal of Fluid Mechanics 96, 65–80.

Baines, P.G., 1982. On internal tide generation models. Deep-Sea Research 29, 307–338.

Balmforth, N.J., Ierley, G.R., Young, W.R., 2002. Tidal conversion by subcritical topography. Journal of Physical Oceanography 32 (10), 2900–2914.

Benjamin, T.B., 1968. Gravity currents and related phenomena. Journal of Fluid Mechanics 31, 209–248.

Beres, J.H., Alexander, M.J., Holton, J.R., 2002. Effects of tropospheric wind shear on the spectrum of convectively generated gravity waves. Journal of Atmospheric Sciences 59, 1805–1824.

Britter, R.E., Simpson, J.E., 1981. A note on the structure of the head of an intrusive gravity current. Journal of Fluid Mechanics 112, 459–466.

Cornillon, P., Stramma, L., Price, J.F., 1987. Satellite measurements of sea surface cooling during Hurricane Gloria. Nature 326, 373–375.

Dalziel, S.B., 1992. Decay of rotating turbulence: some particle tracking experiments. Applied Scientific Research 49, 217–244.

Dalziel, S.B., Hughes, G.O., Sutherland, B.R., 2000. Whole field density measurements. Experiments in Fluids 28, 322–335.

Delisi, D.P., Orlanski, I., 1975. On the role of density jumps in the reflexion and breaking of internal gravity waves. Journal of Fluid Mechanics 69, 445–464.

Dillon, T.M., Moum, J.N., Chereskin, T.K., Caldwell, D.R., 1989. Zonal momentum balance at the equator. Journal of Physical Oceanography 19, 561–570.

Dohan, K., Sutherland, B.R., 2003. Internal waves generated from a turbulent mixed region. Physics of Fluids 15, 488–498.

Faust, K.M., Plate, E.J., 1984. Experimental investigation of intrusive gravity currents entering stably stratified fluids. Journal of Hydraulic Research 22 (5), 315–325.

Flynn, M.R., Sutherland, B.R., 2004. Intrusive gravity currents and internal wave generation in stratified fluid. Journal of Fluid Mechanics 514, 355–383.

Garrett, C., St. Laurent, L.C., 2002. Aspects of deep ocean mixing. Journal of Oceanography 58, 11–24.

Gerkema, T., 2001. Internal and interfacial tides: beam scattering and local generation of solitary waves. Journal of Marine Research 59, 227–255.

Hebert, D., Moum, J.N., Paulson, C.A., Caldwell, D.R., Chereskin, T.K., McPhaden, M.J., 1991. The role of the turbulent stress divergence in the equatorial Pacific zonal momentum balance. Journal of Geophysical Research 96 (C4), 7127–7136.

Holyer, J.Y., Huppert, H.E., 1980. Gravity currents entering a two-layer fluid. Journal of Fluid Mechanics 100, 739–767.

Hughes, G.O., 1996. Aspects of mixing in stratified flows. Ph.D. Thesis, University of Cambridge.

Ledwell, J.R., Montgomery, E., Polzin, K., St. Laurent, L.C., Schmitt, R., Toole, J., 2000. Evidence for enhanced mixing over rough topography in the abyssal ocean. Nature 403, 179–182.

Lien, R.-C., McPhaden, M.J., Gregg, M.C., 1996. High-frequency internal waves at 0°, 140° W and their possible relationship to deep-cycle turbulence. Journal of Physical Oceanography 26, 581–600.

Lowe, R.J., Linden, P.F., Rottman, J.W., 2002. A laboratory study of the velocity structure in an intrusive gravity current. Journal of Fluid Mechanics 456, 33–48.

Manins, P.C., 1976. Intrusion into a stratified media. Journal of Fluid Mechanics 74, 547–560.

Maxworthy, T., Leilich, J., Simpson, J., Meiburg, E.H., 2002. The propagation of a gravity current in a linearly stratified fluid. Journal of Fluid Mechanics 453, 371–394.

Mehta, A., Sutherland, B.R., Kyba, P.J., 2002. Interfacial gravity currents: Part II—wave excitation. Physics of Fluids 14, 3558–3569.

Michaelian, M.E., Maxworthy, T., Redekopp, L.G., 2002. The coupling between turbulent penetrative convection and internal waves. European Journal of Mechanics B 21, 1–28.

Moum, J.N., Hebert, D., Paulson, C.A., Caldwell, D.R., 1992. Turbulence and internal waves at the equator. Part 1: Statistics from towed thermistors and a microstructure profiler. Journal of Physical Oceanography 22, 1330–1345.

Oster, G., 1965. Density gradients. Scientific American 213, 70.

Price, J.F., Sanford, T.B., Forristall, G.Z., 1994. Forced stage response to a moving hurricane. Journal of Physical Oceanography 24, 233–260.

Rooij, F. de, 1999. Sedimenting particle-laden flows in confined geometries. Ph.D. Thesis, University of Cambridge.

Schooley, A.H., Hughes, B.A., 1972. An experimental and theoretical study of internal waves generated by the collapse of a two-dimensional mixed region in a density gradient. Journal of Fluid Mechanics 51, 159–175.

Skyllingstad, E.D., Denbo, D.W., 1994. The role of internal gravity waves in the equatorial current system. Journal of Physical Oceanography 24, 2093–2110.

Smith, S.G.L., Young, W.R., 2002. Conversion of the barotropic tide. Journal of Physical Oceanography 32, 1554–1566.

Smyth, W.D., Moum, J.N., 2002. Shear instability and gravity wave saturation in an asymmetrically stratified jet. Dynamics of Atmospheres and Oceans 35 (3), 265–294.

Smyth, W.D., Peltier, W.R., 1991. Instability and transition in finite-amplitude Kelvin–Helmholtz and Holmboe waves. Journal of Fluid Mechanics 228, 387–415.

St. Laurent, L.C., Toole, J.M., Schmitt, R.W., 2001. Buoyancy forcing by turbulence above rough topography in the abyssal Brazil Basin. Journal of Physical Oceanography 31, 3476–3495.

Sutherland, B.R., 1996. The dynamic excitation of internal gravity waves in the equatorial oceans. Journal of Physical Oceanography 26, 3214–3235.

Sutherland, B.R., 2002. Large-amplitude internal wave generation in the lee of step-shaped topography. Geophysical Research Letters 29(16), art. no 1769.

Sutherland, B.R., Linden, P.F., 1998. Internal wave generation by flow over a thin barrier. Journal of Fluid Mechanics 377, 223–252.

Sutherland, B.R., Peltier, W.R., 1994. Turbulence transition and internal wave generation in density stratified jets. Physics of Fluids A 6, 1267–1284.

Sutherland, B.R., Dalziel, S.B., Hughes, G.O., Linden, P.F., 1999. Visualisation and measurement of internal waves by "synthetic schlieren". Part 1: Vertically oscillating cylinder. Journal of Fluid Mechanics 390, 93–126.

Thorpe, S.A., 1998. Nonlinear reflection of internal waves at the density discontinuity at the base of a mixed layer. Journal of Physical Oceanography 28, 1853–1860.

Ungarish, M., Huppert, H.E., 2002. On gravity currents propagating at the base of a stratified fluid. Journal of Fluid Mechanics 458, 283–301.

Wijesekera, H.W., Dillon, T.M., 1991. Internal waves and mixing in the upper equatorial Pacific ocean. Journal of Geophysical Research 96, 7115–7126.

Wu, J., 1969. Mixed region collapse with internal wave generation in a density stratified medium. Journal of Fluid Mechanics 35, 531–544.

ELSEVIER

Deep-Sea Research II 51 (2004) 2905–2917

DEEP-SEA RESEARCH
PART II

www.elsevier.com/locate/dsr2

Internal waves in a three-layer bubbly waveguide

R.H.J. Grimshaw, K.R. Khusnutdinova*,[1]

Department of Mathematical Sciences, Loughborough University, Loughborough LE11 3TU, UK

Abstract

In many oceanic situations, clouds, or layers of bubbles are present which may influence the larger-scale dynamics. A two-dimensional model of a diluted locally monodisperse mixture of an incompressible fluid with gas bubbles is used to model and study such situations. The problem is simplified by ignoring the horizontal variability in the structure of the bubble layer and by focusing instead on the vertical variability. We consider waves propagating horizontally in the oceanic waveguide, assuming that bubbles are confined to a thin upper layer of the otherwise homogeneous upper ocean, and using a three layer model to represent this situation. The presence of the depth-dependent distribution of bubbles introduces an effective stratification and considerably changes the value of the buoyancy frequency N_1 in the absence of bubbles, replacing it with an effective value N, where $N^2 = N_1^2 + g\alpha_{g0}[(\ln n_0)_z + 3\alpha_{g0}(\ln R_0)_z]/(1-\alpha_{g0})$ (here α_{g0} is the void fraction, n_0 is the number density, and R_0 is the radius of bubbles in the basic state).This leads to the possibility of existence of the internal waves in the otherwise homogeneous upper mixed layer. Also the presence of the bubbly layer causes significant changes to the dispersion relation for the usual internal waves.
© 2004 Elsevier Ltd. All rights reserved.

1. Introduction

There is considerable observational and experimental evidence about the existence of bubble clouds, and, at higher wind speeds, of bubble layers in the upper ocean. The bubble distributions in the ocean surface layer have been studied in many different contexts: in connection with their role in air–sea gas transfer and as a measure of turbulence (Kolovaev, 1976; Medwin, 1977; Johnson and Cooke, 1979; Thorpe, 1982; Thorpe and Hall, 1983; Thorpe, 1984a; Woolf and Thorpe, 1991; Thorpe, 1995; Melville, 1996; Farmer et al., 1999; Terrill and Melville, 1999; Dean and Stokes, 1999; Garrett et al., 2000), and also in connection with their influence on ambient noise (Medwin, 1977; Urick, 1983; Farmer and Lemon, 1984; Farmer et al., 1997; Farmer and Vagle, 1989; Buckingham, 1997). Zedel and Farmer (1991) have observed the evolution of bubble layers during a storm.

The observed bubble distributions vary both as a function of time and depth and depend significantly on the wind speed. According to

*Corresponding author.
E-mail addresses: r.h.j.grimshaw@lboro.ac.uk (R.H.J. Grimshaw), k.khusnutdinova@lboro.ac.uk (K.R. Khusnutdinova).
[1]On leave from: Institute of Mechanics, Ufa Branch of the Russian Academy of Sciences, Karl Marx Str. 6, Ufa 450000, Russia.

0967-0645/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dsr2.2004.09.003

Thorpe (1982), at winds exceeding 6.5 ms^{-1} clouds of bubbles tend to overlap forming a continuous layer of variable thickness. According to Farmer and Vagle (1989), Zedel and Farmer (1991), and Farmer et al. (1999), this highly structured bubble layer extends to a depth of several meters beneath the surface. Bubble concentration profiles have been measured in wind speeds of 11–13 ms^{-1} by Kolovaev (1976), Johnson and Cooke (1979), and Crawford and Farmer (1987). In general, there is a monotonic decrease in void fraction with increasing depth: the void fraction profile either decays exponentially or follows the inverse-square profile (Buckingham, 1997). According to Farmer and Lemon (1984), the bubble concentration decreases exponentially with depth, with an *e*-folding scale of order 1 m which depends on the wind speed. A 20-fold increase in bubble concentration was reported by Johnson and Cooke (1979) between the wind speed range 8.6–10.7 and 11.8–13.9 ms^{-1}. At wind speeds between 10 and 14 ms^{-1} the void fraction at the surface, averaged over a considerable period of time, is of order 10^{-7}–10^{-6} (see, e.g., Buckingham, 1997 and the references therein; Farmer et al., 1999; Garrett et al., 2000). It can be larger than that for higher winds, or briefly so after the breaking of a wave. For example, at a wind speed of 15 ms^{-1} in observations by Terrill and Melville (1999), the void fractions approached levels of 10^{-4} at 0.69 m depth, while the highest densities measured at 4.1 m were approximately 5×10^{-6}. The mean void fraction at a depth of 0.69 m calculated for a 20 min record was of order 10^{-5}, while at 4.1 m it was of order 10^{-6}. The distributions peak at all depths, and they decrease rapidly on either side of the peak and also rapidly with depth (see, e.g., Thorpe, 1984b; Farmer et al., 1999 and the references there). The observations have shown that bubbles with radii of approximately 100 μm contribute most to the total void fraction. In the observations reported by Farmer et al. (1999), in winds of approximately 12 ms^{-1} the peak radius declined from 130 μm at 1.3 m to 70 μm at 5.5 m. In the observations made by Johnson and Cooke (1979) in winds of 11–13 ms^{-1} the peak was at the radius of about 50 μm at all depths. According to Scott (1975), Thorpe (1984b) and Bullock et al. (2001) changes in salinity may have a profound effect on the bubble size distributions: far more small bubbles are formed in saline water than in fresh water.

As in our previous paper (Grimshaw and Khusnutdinova, 2004), we follow an approach developed in the mechanics of multiphase media (see Nigmatulin, 1991; Nakoryakov, et al.1983; Kutateladze and Nakoryakov, 1984; Prosperetti, 1994; Brennen, 1995 and references therein) to study the properties of the linear internal waves in a continuously and uniformly stratified fluid in the presence of bubbles. In the ocean surface layer bubbles occur mainly due to the breaking of surface waves, and evolve under the influence of buoyancy, turbulence, currents, dissolution and bubble break-up. Here, we consider the later stages of the fully developed sea, when conditions can be assumed to be quasi-steady. Following Thorpe (1982), we assume that "the tendency of bubbles to rise is balanced by turbulent diffusion" and take the view that there is a steady-state bubble distribution in place, which is then perturbed. The processes of bubble formation, destruction and dissolution are not taken into account. Shear flow can be taken into account in our model (by choosing a more complicated basic state). However, here we follow the traditional approach towards studying internal waves and start from a simpler model, with no shear flow. Our present paper is organized as follows. In Section 2 we formulate the mathematical model describing the two-dimensional motion of a diluted locally monodisperse mixture of incompressible fluid containing small gas bubbles. In Section 3 we briefly discuss the dispersion relation for the plane linear waves. In Section 4 we present some results for waves which are constrained to propagate horizontally in an oceanic waveguide, the upper part of which contains the depth-dependent distribution of bubbles. In our previous paper (Grimshaw and Khusnutdinova, 2004) we used a two-layer model to represent this situation. Here, we make a more realistic assumption that bubbles are confined to a thin upper layer of the homogeneous upper ocean, and use a three layer model to obtain the dispersion relation, accounting for the mismatch between the depth of the bubble layer and the depth of the upper mixed layer. This

also allows us to define the region of validity of our previously obtained results more precisely.

It should be mentioned, that, in recent years, bubbles have been intensively studied in many different situations. For instance, surface waves in a bubbly fluid, and waves on the interface between bubbly and ideal fluids in the absence of stratification were discussed by Akhatov (1989). Acoustic generation in a bubble layer has been studied in Druzhinin et al. (1996), Ostrovsky et al. (1998), and Karpov et al. (2003). There also has been significant interest in studying the dynamics of clouds of bubbles in connection with cavitation problems (d'Agostino and Brennen, 1989; Kumar and Brennen, 1993; Brennen, 1998; Reisman et al., 1998; see also references in Reisman et al., 1998). The relationship between internal waves and bubbles has been observed by Thorpe et al. (1987).

2. Mathematical model

We use a model describing the two-dimensional motion of a diluted locally monodisperse mixture of incompressible fluid containing small gas bubbles, preserving their mass and spherical form. Although the real oceanic flow is polydisperse, the use of a locally monodisperse model as an approximation to the real oceanic situation is based on the existence of a peak in the bubble distributions at all depths, as discussed in the Introduction. The distances between the bubbles are supposed to be big enough to prevent collisions, and so the interaction between bubbles is due only to pressure changes. The processes of bubble formation and destruction are not taken into account. The fluid phase is assumed to be incompressible, following the traditional approach towards studying internal waves (in the absence of bubbles). This is also justified by the great compressibility of bubbles compared to that of the surrounding liquid, and their small dimensions. The void fraction α_g is supposed to be small enough, $\alpha_g \ll 1$, so that the mass of the gas is neglected compared to the mass of the fluid. Thus, the mixture can be considered as a medium with the density approximately equal to $\rho_l(1-\alpha_g)$. Here, dissipative mechanisms are not taken into account (for a model accounting for the dissipation see Grimshaw and Khusnutdinova, 2004).

Under these assumptions the set of equations describing a two-dimensional flow takes the form

$$\rho \frac{\mathrm{d}u}{\mathrm{d}t} + p_x = 0, \quad \rho \frac{\mathrm{d}w}{\mathrm{d}t} + p_z + \rho g = 0, \tag{1}$$

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} + \rho(u_x + w_z) = 0, \quad \frac{\mathrm{d}n}{\mathrm{d}t} + n(u_x + w_z) = 0, \tag{2}$$

$$\rho_l \left(R \frac{\mathrm{d}^2 R}{\mathrm{d}t^2} + \frac{3}{2} \left(\frac{\mathrm{d}R}{\mathrm{d}t} \right)^2 \right) = p_g - p - 2\frac{\sum}{R}, \tag{3}$$

$$\frac{\mathrm{d}\rho_l}{\mathrm{d}t} = 0, \quad \rho = \rho_l(1-\alpha_g), \quad \alpha_g = \frac{4}{3}\pi n R^3, \tag{4}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(p_g R^{3\kappa}) = 0. \tag{5}$$

Here p, ρ, u, w are pressure, density and velocity components of the mixture; ρ_l is the density of a pure fluid; p_g, α_g, R, n are pressure, void fraction, radius and number density of bubbles; κ is a polytropic exponent of a gas ($1 \leqslant \kappa \leqslant \gamma$, where γ is the adiabatic exponent); $\mathrm{d}/\mathrm{d}t = \partial/\partial t + u\,\partial/\partial x + w\,\partial/\partial z$ is the material derivative with respect to time; and $\sum$ is the coefficient of surface tension. Here, the basic hydrodynamic equations for the mixture are closed by the Rayleigh–Plesset equation (3) for the oscillations of a single bubble, where the pressure at infinity is replaced by the pressure of the mixture (see Rayleigh, 1917; Iordansky, 1960; Kogarko, 1961; van Wijngaarden, 1968; Plesset and Prosperetti, 1977; Nigmatulin, 1991). Eq. (5) follows from the conservation laws for entropy of the gas, and the mass of the bubble, that is,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\ln p_g \rho_g^{-\kappa}) = 0 \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}t}(\rho_g R^3) = 0,$$

respectively, where ρ_g is the density of the gas. The details of derivation of Eq. (3), as well as discussion of the above model as a whole can be found, for example, in Nigmatulin (1991).

It is convenient to rewrite these equations in a different form, using (4). Indeed, from (4) we

find that

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\rho_1 \frac{\mathrm{d}\alpha_g}{\mathrm{d}t} = -\frac{\rho}{1-\alpha_g}\frac{\mathrm{d}\alpha_g}{\mathrm{d}t}.$$

From this, using the last relation in (4), we find that

$$\frac{1}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\frac{1}{1-\alpha_g}\left(\alpha_g \frac{1}{n}\frac{\mathrm{d}n}{\mathrm{d}t} + 3\alpha_g \frac{1}{R}\frac{\mathrm{d}R}{\mathrm{d}t}\right), \tag{6}$$

while, from (2),

$$\frac{1}{n}\frac{\mathrm{d}n}{\mathrm{d}t} = \frac{1}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}t}. \tag{7}$$

Finally, substituting (7) into (6), we obtain

$$\frac{1}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}t} = -3\alpha_g \frac{1}{R}\frac{\mathrm{d}R}{\mathrm{d}t}. \tag{8}$$

Then, using (8), we can rewrite the governing equations (1)–(5) as follows:

$$\rho\frac{\mathrm{d}u}{\mathrm{d}t} + p_x = 0, \quad \rho\frac{\mathrm{d}w}{\mathrm{d}t} + p_z + \rho g = 0, \tag{9}$$

$$\frac{1}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{1}{n}\frac{\mathrm{d}n}{\mathrm{d}t} = -3\alpha_{\mathrm{g}}\frac{1}{R}\frac{\mathrm{d}R}{\mathrm{d}t}, \tag{10}$$

$$u_x + w_z = 3\alpha_{\mathrm{g}}\frac{1}{R}\frac{\mathrm{d}R}{\mathrm{d}t}, \tag{11}$$

$$\rho\left(R\frac{\mathrm{d}^2R}{\mathrm{d}t^2} + \frac{3}{2}\left(\frac{\mathrm{d}R}{\mathrm{d}t}\right)^2\right) = (1-\alpha_{\mathrm{g}})\left(p_{\mathrm{g}} - p - 2\frac{\sum}{R}\right), \tag{12}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(p_{\mathrm{g}}R^{3\kappa}) = 0. \tag{13}$$

When written in this form, the governing equations can be more conveniently viewed as an extension, due to the bubbles, of the usual equations for the internal waves in incompressible fluid, which simplifies their analysis. Thus, the model we are working with consists of seven equations (9)–(13) in seven unknowns $p, \rho, u, w, R, n, p_{\mathrm{g}}$, with $\alpha_{\mathrm{g}} = 4\pi nR^3/3$. In Section 4 we will supplement these equations by boundary conditions at the ocean bottom and at the free surface in order to discuss horizontally propagating waves.

We will suppose that in the basic state the mixture has density $\rho_0(z)$, a corresponding pressure $p_0(z)$ (satisfying $p_{0z} = -g\rho_0$), and that there is no shear flow, $u_0 = w_0 = 0$. We also have $p_{g0}(z) = p_0(z) + 2\sum/R_0$ and $\alpha_{g0} = 4\pi n_0 R_0^3/3$. Here, we take into consideration only the vertical variability of the bubble layer and do not consider the horizontal spatial variations of its thickness, which in reality is a significant feature of the upper ocean. However, we think that this is a correct idealization of the real situation, allowing us to obtain a simplified problem and to focus on the effects due to the vertical variability of the bubble layer.

Using the Bousinesq approximation, the linearized equations describing deviations from the basic state take the form

$$\rho_0\tilde{u}_t + \tilde{p}_x = 0, \tag{14}$$

$$\rho_0\tilde{w}_t + \tilde{p}_z + g\tilde{\rho} = 0, \tag{15}$$

$$\tilde{\rho}_t + \rho_{0z}\tilde{w} = -3\frac{\alpha_{g0}\rho_0}{R_0}(\tilde{R}_t + R_{0z}\tilde{w}), \tag{16}$$

$$\tilde{u}_x + \tilde{w}_z = 3\frac{\alpha_{g0}}{R_0}(\tilde{R}_t + R_{0z}\tilde{w}), \tag{17}$$

$$\rho_0 R_0(\tilde{R}_{tt} + R_{0z}\tilde{w}_t) = (1-\alpha_{g0})\left(\tilde{p}_{\mathrm{g}} - \tilde{p} + 2\frac{\sum}{R_0^2}\tilde{R}\right), \tag{18}$$

$$\tilde{p}_{gt} + \frac{3\kappa p_{g0}}{R_0}\tilde{R}_t + \frac{(p_{g0}R_0^{3\kappa})_z}{R_0^{3\kappa}}\tilde{w} = 0, \tag{19}$$

$$\tilde{n}_t + n_{0z}\tilde{w} = -3\frac{\alpha_{g0}n_0}{R_0}(\tilde{R}_t + R_{0z}\tilde{w}). \tag{20}$$

In the sequel, the *tilde* superscripts are omitted. Note that n appears only in the last equation, and so does not influence the dispersion relation.

It is convenient to work with the equations involving only the vertical velocity w, and $s = R_t$, the radial velocity of the bubbles (which in this model is assumed to be described by one function for all bubbles). The derivation of these equations is given below. The horizontal velocity u can be eliminated by taking the time derivative of (17) and using the horizontal momentum

equation (14), which leads to

$$\frac{1}{\rho_0} p_{xx} = w_{zt} - \frac{3\alpha_{g0}}{R_0} (R_{tt} + R_{0z} w_t). \tag{21}$$

Then on taking the time derivative of the vertical momentum equation (15) and eliminating ρ by virtue of (16), we get

$$\frac{1}{\rho_0} p_{zt} = -w_{tt} - N^2 w + \frac{3\alpha_{g0} g}{R_0} R_t, \tag{22}$$

where

$$N^2 = -g\left(\frac{1}{\rho_0}\frac{\mathrm{d}\rho_0}{\mathrm{d}z} + \frac{3\alpha_{g0}}{R_0}\frac{\mathrm{d}R_0}{\mathrm{d}z}\right). \tag{23}$$

The pressure p can be eliminated by taking the second derivative of (22) with respect to x, and using (21). In the Boussinesq approximation we obtain

$$\frac{\partial^2}{\partial t^2} \Delta w + N^2 w_{xx} - [b(s_{tt} + R_{0z} w_{tt})]_z - g b s_{xx} = 0, \tag{24}$$

where $\Delta \equiv \partial^2/\partial x^2 + \partial^2/\partial z^2$ is the two-dimensional Laplacian, and $b = 3\alpha_{g0}/R_0$. Here *the effective buoyancy frequency* N (see (23)) is given by

$$N^2 = N_1^2 + g\left(\frac{\alpha_{g0z}}{1-\alpha_{g0}} - 3\alpha_{g0}\frac{R_{0z}}{R_0}\right)$$
$$= N_1^2 + \frac{g\alpha_{g0}}{1-\alpha_{g0}}[(\ln n_0)_z + 3\alpha_{g0}(\ln R_0)_z],$$

where N_1 is the usual buoyancy frequency of the liquid (in the absence of bubbles). Here, we have used the relation $\alpha_{g0} = 4\pi n_0 R_0^3/3$. When the void fraction $\alpha_{g0} \ll 1$, the term involving the number density gradient $(\ln n_0)_z$ dominates that due to the gradient of the bubble radius $(\ln R_0)_z$. Since N_1^2 in the upper mixed layer is usually very small and $(\ln n_0)_z$ is of order 1, then even for α_{g0} of order 10^{-8}, the effective buoyancy frequency N significantly differs from the value of the buoyancy frequency N_1 in the absence of any bubbles. For larger void fractions of bubbles (for example, for $\alpha_{g0} \sim 10^{-6} - 10^{-5}$, which is not a rare occurrence in the ocean) we can get anomalous high values of the effective buoyancy frequency in the upper ocean, which leads to the possibility of existence of subsurface internal waves similar to the usual internal waves in deep water (see also the end of Section 4). Note that this significant influence of bubbles on internal waves manifests itself already in the linear approximation.

A second equation can be derived by differentiating the linearized Rayleigh–Plesset equation (18) twice by x and once by t, and then using (19) and (21). This leads to

$$s_{ttxx} + \omega_*^2(1-\alpha_{g0})s_{xx} - a s_{tt}$$
$$+ \frac{a}{b}\left(w_{ttz} + \frac{(p_{g0}R_0^{3\kappa})_z}{\rho_0 R_0^{3\kappa}} w_{xx}\right)$$
$$+ R_{0z}(w_{xx} - a w)_{tt} = 0, \tag{25}$$

where

$$\omega_*^2 = \frac{3\kappa p_0}{R_0^2 \rho_0} + \frac{2\sigma(3\kappa - 1)}{R_0^3} \tag{26}$$

is the natural frequency of bubble oscillations, $\sigma = \sum/\rho_0$, and $a = 3\alpha_{g0}(1-\alpha_{g0})/R_0^2$. Note that if $\sigma = 0$, then $\omega_*^2 = 3\kappa p_0/R_0^2\rho_0$ is a Minnaert frequency (Minnaert, 1933).

We now assume that the radius R_0 of bubbles in the basic state is independent of the depth. This approximation is based on the observations discussed in the Introduction, from which we argue that the changes in the number of bubbles with depth are much more crucial. Then, introducing the new variable (the depth-dependent scaling of the radial velocity of the bubble),

$$c = \alpha_{g0}(z)s$$

Eqs. (24) and (25) can be brought to the form

$$\frac{\partial^2}{\partial t^2} \Delta w + N^2 w_{xx} - \delta(c_{ttz} + g c_{xx}) = 0,$$

$$c_{ttxx} + \omega_*^2(1-\alpha_{g0})c_{xx} - a c_{tt}$$
$$+ \frac{a}{\delta}(w_{ttz} - g w_{xx}) = 0. \tag{27}$$

Here $\delta = 3/R_0$. We shall next simplify the analysis by assuming that the effective buoyancy frequency N (23) is constant, at least in each of the several layers (see Section 4). This then gives us $\rho_0 = \rho_a \exp(-N^2 z/g)$ and $p_0 = p_a + g^2(\rho_0 - \rho_a)/N^2$, where p_a and ρ_a are constants, which may be taken as the atmospheric pressure and the density at the free surface, respectively. However, even in the case of constant N, ω_* (26) and the

parameter α_{g0} in the coefficients of (27) depend on z. In our present analysis we consider a simplified model by supposing that the values of ω_* and α_{g0} in the coefficients of (27) are constant and equal to their average value over the depth of the bubble layer. Thus, in this simplified model, the depth-dependence of the actual distribution of bubbles is taken into account only by introducing the effective buoyancy frequency N. We shall also make the usual assumption that σ is a constant.

3. Dispersion relation

The first step is to obtain the dispersion relation, by seeking solutions in the form of propagating waves. Considering two-dimensional plane waves

$$\begin{pmatrix} w \\ c \end{pmatrix} = \begin{pmatrix} W \\ C \end{pmatrix} e^{i(kx+lz-\omega t)} + c.c. \tag{28}$$

We obtain the dispersion relation

$$a_1\,\omega^4 - a_2\,\omega^2 + a_3 = 0, \tag{29}$$

where the coefficients are given by

$$a_1 = k^2 + l^2 + a,$$
$$a_2 = N^2(k^2 + a) + \omega_*^2(1 - \alpha_{g0})(k^2 + l^2),$$
$$a_3 = (\omega_*^2 N^2(1 - \alpha_{g0}) - ag^2)k^2.$$

From (29) we obtain

$$\omega_\pm^2 = \frac{a_2 \pm \sqrt{D}}{2a_1}, \quad D = a_2^2 - 4a_1a_3. \tag{30}$$

The expression for the discriminant D can be rewritten as

$$D = (\omega_*^2(1 - \alpha_{g0})(k^2 + l^2) - N^2(k^2 + a))^2 + 4ag^2k^2(k^2 + l^2 + a) + 4\omega_*^2(1 - \alpha_{g0})N^2al^2,$$

showing that $D > 0$. Next, we observe that if $a_3 < 0$, then $\omega_-^2 < 0$ implying the existence of a temporal instability. However, since $a_3 < 0$ is equivalent to $\omega_*^2 R_0^2 N^2 < 3\alpha_{g0}g^2$, it is unlikely to be satisfied under the hypotheses we have used in deriving this set of equations. Therefore, we shall always assume that $a_3 > 0$. In this case, $D < a_2^2$, and (30) shows that the dispersion relation has two branches, ω_+ and ω_-, characterized by the following properties:

Let the vertical wavenumber l be fixed. Then, as $k \to \infty$,

$$\omega_\pm^2 \to \omega_{\pm(\infty)}^2 = \tfrac{1}{2}\Big(\omega_*^2(1 - \alpha_{g0}) + N^2 \pm \sqrt{(\omega_*^2(1 - \alpha_{g0}) - N^2)^2 + 4ag^2}\Big).$$

Taking into account that for typical values of parameters $4ag^2 \ll (\omega_*^2(1 - \alpha_{g0}) - N^2)^2$ and $N^2 \ll \omega_*^2(1 - \alpha_{g0})$, we get

$$\omega_{+(\infty)}^2 \approx \omega_*^2(1 - \alpha_{g0}) + \frac{3\alpha_{g0}g^2}{\omega_*^2 R_0^2},$$

$$\omega_{-(\infty)}^2 \approx N^2 - \frac{3\alpha_{g0}g^2}{\omega_*^2 R_0^2}.$$

On the other hand, as $k \to 0$,

$$\omega_+^2 \to \omega_{+(0)}^2 = \frac{\omega_*^2(1 - \alpha_{g0})l^2 + aN^2}{l^2 + a},$$

$$\omega_-^2 \to \omega_{-(0)}^2 = 0.$$

When k is fixed and $l \to \infty$, then

$$\omega_+^2 \to \omega_*^2(1 - \alpha_{g0}), \quad \omega_-^2 \to 0.$$

If $\alpha_{g0} \to 0$, then

$$\omega_+^2 = \omega_*^2 \quad \text{and} \quad \omega_-^2 = \frac{N^2k^2}{k^2 + l^2}.$$

Thus, ω_+ corresponds to the natural frequency of bubble oscillations, while ω_- gives us the dispersion relation for the internal waves in a pure stratified fluid. When $\alpha_{g0} > 0$, corresponding to the presence of bubbles, we see that the frequency of an internal wave lies in the band $0 \leqslant \omega < \omega_{-(\infty)}$, and, for a fixed wavenumber (k, l), significantly differs from the value in the absence of bubbles (for the uniform distribution of bubbles it is reduced). For the depth-dependent distribution of bubbles, the frequency of an internal waves is reduced from the value of the effective buoyancy frequency (23). For example, considering the typical oceanic situation, let the buoyancy frequency of the liquid be $N_1 = 10^{-3}\,s^{-1}$ and the mean void fraction of bubbles be $\alpha_{g0} = 5 \times 10^{-6}$. Supposing that $(\ln n_0)_z \approx 1$, we estimate the value of the effective buoyancy frequency as $N \approx 7.07 \times 10^{-3}\,s^{-1}$. For $\omega_* = 2.26 \times 10^5\,s^{-1}$ and

$R_0 = 100\,\mu m$, the frequency of an internal wave is bounded by a slightly reduced value of $\omega_{-(\infty)} \approx 6.87\times10^{-3}\,s^{-1}$. This value significantly differs from the value $N_1 = 10^{-3}\,s^{-1}$ in the absence of bubbles. The "bubble" wave frequency lies in the band $\omega_{+(0)} \leqslant \omega < \omega_{+(\infty)}$. Since $\omega_{+(0)} \geqslant N > \omega_{-(\infty)}$, there is a gap in the frequency spectrum, and this gap contains the effective buoyancy frequency N (23). In particular, the limiting frequency for an internal wave is $\omega_{-(\infty)}$ which is less than N.

The group velocity for each branch is $(\partial\omega/\partial k, \partial\omega/\partial l)$, and can be written in the form:

$$\frac{\partial\omega_\pm}{\partial k} = \frac{k}{\omega_\pm}\frac{(\omega_\pm^2 - \omega_{+(\infty)}^2)(\omega_\pm^2 - \omega_{-(\infty)}^2)}{\mp\sqrt{D}},$$
$$\frac{\partial\omega_\pm}{\partial l} = \omega_\pm l\frac{\omega_\pm^2 - \omega_*^2(1-\alpha_{g0})}{\mp\sqrt{D}}. \quad (31)$$

It follows that (assuming without loss of generality that $\omega_\pm > 0$) for the internal wave, the horizontal component of group velocity is always positive, and the vertical component is always negative (as for the case when there are no bubbles). For the bubble mode, both the horizontal and the vertical components of group velocity are positive, except for short waves ($k \to \infty$) for which $\omega_+^2 > \omega_*^2(1-\alpha_{g0})$ when the vertical component is negative. Further, it can be shown that for the internal wave, the horizontal component of group velocity is reduced from that in a pure fluid with the buoyancy frequency equal to the effective value (23), while the absolute value of the vertical component is increased, except for very long waves for which $k, l \to 0$.

If $\alpha_{g0} \to 0$, then (31) yields

$$\left(\frac{\partial\omega_-}{\partial k}, \frac{\partial\omega_-}{\partial l}\right) \to \frac{l\omega_-}{k(k^2+l^2)}(l, -k),$$

which is the group velocity for internal waves in a pure stratified fluid, and is normal to the phase velocity. Let us note, that in the presence of bubbles the group velocity is no longer normal to the phase velocity, since, as can be seen from (17), the particle motion is not normal to the wave-number vector. Further, for $l = 0$, the group velocity of internal waves in a pure fluid is zero, but in the presence of bubbles it has a non-zero horizontal component.

Note, that if we put $g = 0$ in our formulas, the dispersion relation collapses to $\omega^2 = 0$ (the remnant of the internal wave branch) and a single non-trivial branch:

$$\omega^2 = \frac{\omega_*^2(k^2+l^2)}{k^2+l^2+a}. \quad (32)$$

For long waves (i.e. $k^2 + l^2 \ll a$) (32) yields the well-known formula for the sound velocity in a gas–liquid mixture (e.g., Naugolnykh and Ostrovsky, 1998):

$$\omega^2 \approx c_0^2(k^2+l^2), \quad c_0^2 = \frac{\kappa p_0}{\rho_0\alpha_{g0}}.$$

Thus, returning to the dispersion relation (30) we can say that the bubble mode (the upper branch of the dispersion relation) can be interpreted as modified acoustic waves, while the lower branch represents modified internal waves (but note that for very long waves there is an exchange of identity between the branches). Although the liquid is assumed to be incompressible, the presence of bubbles gives the mixture some compressibility.

4. Waves in a three-layer bubbly waveguide

We now consider the oceanic waveguide, and suppose that the fluid is bounded below by a flat rigid bottom, and above by a free surface. Here, we assume that there is a thin upper layer of the homogeneous upper ocean which has a bubble distribution, and there are no bubbles in the rest of the upper mixed layer and in the deep ocean. We

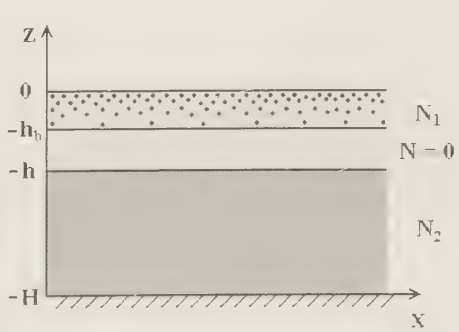


Fig. 1. The three-layer bubbly waveguide.

use a three-layer model, shown in Fig. 1, to represent this situation.

The upper bubbly layer has an undisturbed constant depth h_b, a constant effective buoyancy frequency N_1 and contains bubbles. This bubble layer is a part of the otherwise homogeneous upper mixed layer, which has an undisturbed constant depth h. The lower layer is without bubbles, and has a constant buoyancy frequency N_2. It is important to note, that although we assume the upper layer to be homogeneous in the absence of bubbles, the effective buoyancy frequency N_1 can be comparable (or even bigger) than the buoyancy frequency N_2 in the lower stratified pure fluid layer. The density $\rho_0(z)$ is assumed to be continuous at both interfaces: $z=-h_b$ and $z=-h$. In the sequel the subscript 1 corresponds to the upper bubble layer, the subscript 2 to the middle homogeneous pure liquid layer, and the subscript 3 to the lower stratified pure liquid layer. The linearized equations of motion are again (14)–(20). The bottom boundary condition is

$$w_3|_{z=-H}=0, \tag{33}$$

where H is the undisturbed fluid depth. In the Boussinesq approximation for internal waves, we may replace the free-surface boundary conditions with the so-called "rigid-lid" condition (e.g., Gill, 1982):

$$w_1|_{z=0}=0. \tag{34}$$

The linearized conditions at the first interface (between the upper bubble layer and the middle layer) are

$$w_1|_{z=-h_b}=w_2|_{z=-h_b}=\eta_t, \tag{35}$$

$$p_1|_{z=-h_b}=p_2|_{z=-h_b}, \tag{36}$$

expressing continuity of the vertical displacement and the pressure (in (36), we have taken into account that there is no density jump at the interface). Here, $\eta(x,t)$ is the first interface displacement. Similarly, the linearized conditions at the second interface (between the middle layer and the lower layer) are

$$w_2|_{z=-h}=w_3|_{z=-h}=\zeta_t, \tag{37}$$

$$p_2|_{z=-h}=p_3|_{z=-h}, \tag{38}$$

where $\zeta(x,t)$ is the second interface displacement.

In the lower stratified layer (pure fluid without bubbles) we have the equation

$$\frac{\partial^2}{\partial t^2}\Delta w_3+N_2^2 w_{3xx}=0. \tag{39}$$

The solution of (39), which satisfies (33) has the form

$$w_3=A\sin l(z+H)\exp[\mathrm{i}(kx-\omega t)]+c.c., \tag{40}$$

where

$$\omega^2=\frac{N_2^2k^2}{k^2+l^2} \tag{41}$$

is the dispersion relation in a pure stratified fluid. For an internal wave we assume that $l^2>0$ and so then $\omega^2<N_2^2$, although the following analysis remains valid if instead $l^2<0$. In the homogeneous middle layer we have

$$w_2=(B\sinh kz+C\cosh kz)\times\exp[\mathrm{i}(kx-\omega t)]+c.c. \tag{42}$$

In the upper bubble layer we have Eqs. (27) as in Section 3, but with N replaced by N_1. Here we assume that the coefficients ω_*, α_{g0} are replaced by constants, equal to their average value over the depth of the bubble layer. To the leading order in the Boussinesq approximation, we have

$$\omega_*^2\approx\frac{3\kappa}{R_0^2}\left(\frac{p_a}{\rho_a}+\frac{gh_b}{2}\right)+\frac{2\sigma(3\kappa-1)}{R_0^3}.$$

Here p_a is the atmospheric pressure, and ρ_a is the density at the free surface. The solutions of (27) can be found in the form

$$w_1=\phi(z)\exp[\mathrm{i}(kx-\omega t)]+c.c.,$$
$$c=\psi(z)\exp[\mathrm{i}(kx-\omega t)]+c.c. \tag{43}$$

Substituting (43) into (27) we obtain

$$\phi''-\beta^2\phi=0,\quad \psi=\alpha(\omega^2\phi'-gk^2\phi), \tag{44}$$

where

$$\beta^2=\frac{(\omega^2-N_1^2-\alpha\delta g^2k^2)k^2}{\omega^2(1-\alpha\delta\omega^2)},$$
$$\alpha=\frac{a}{\delta[\omega^2(k^2+a)-\omega_*^2(1-\alpha_{g0})k^2]}. \tag{45}$$

It is convenient to assume that $\beta^2 > 0$ which is the natural choice for internal waves when $N_1 < N_2$. However, in fact the following analysis remains valid if $\beta^2 < 0$, and situations when this occurs will be discussed below. Thus Eq. (44) has the solution

$$\phi(z) = D \sinh \beta z,$$

satisfying condition (34).

Next, to apply the boundary conditions at the interfaces we let $\eta(x,t) = \mathrm{i}\tilde{a}\exp[\mathrm{i}(kx - \omega t)] + c.c.$ and $\zeta(x,t) = \mathrm{i}\tilde{b}\exp[\mathrm{i}(kx - \omega t)] + c.c.$ Substituting (43) and (41) into (35), we obtain

$$D = -\frac{\tilde{a}\omega}{\sinh \beta h_b},$$

$$B \sinh kh_b - C \cosh kh_b = -\tilde{a}\omega.$$

Condition (36) can be rewritten in terms of the variables w_1, c, w_2 by differentiating (36) twice with respect to x and using (21) in the bubbly layer, and the similar condition with $\alpha_{g0} = 0$ in the pure fluid. We obtain

$$w_{1_{zt}} - \delta c_t|_{z=-h_b} = w_{2_{zt}}|_{z=-h_b},$$

which leads to

$$(1 - \alpha\delta\omega^2)\beta \coth \beta h_b - k\frac{B \cosh kh_b - C \sinh kh_b}{B \sinh kh_b - C \cosh kh_b} = \alpha\delta g k^2. \tag{46}$$

Now, substituting (42) and (40) into (37), we obtain

$$A = \frac{\tilde{b}\omega}{\sin l(H-h)},$$

$$B \sinh kh - C \cosh kh = -\tilde{b}\omega. \tag{47}$$

As before, we can rewrite (38) in terms of the variables w_2 and w_3:

$$w_{2_{zt}}|_{z=-h} = w_{3_{zt}}|_{z=-h},$$

which gives us

$$k(B \cosh kh - C \sinh kh) = Al \cos l(H-h). \tag{48}$$

Finally, defining B and C from (47) and (48), and substituting the values of B and C into (46), we obtain the dispersion relation,

$$(1 - \alpha\delta\omega^2)\beta \coth \beta h_b - \alpha\delta g k^2 = -l\frac{\cos(H-h)l \cosh(h-h_b)k + k/l \sin(H-h)l \sinh(h-h_b)k}{\sin(H-h)l \cosh(h-h_b)k + l/k \cos(H-h)l \sinh(h-h_b)k}, \tag{49}$$

accounting for the mismatch between the depth of the bubble layer and the depth of the upper mixed layer. Note that $k^2 = \omega^2 l^2/(N_2^2 - \omega^2)$, and, therefore, from (45), we obtain

$$1 - \alpha\delta\omega^2 = \frac{(\omega_*^2(1-\alpha_{g0}) - \omega^2)l^2}{(\omega_*^2(1-\alpha_{g0}) - \omega^2)l^2 - a(N_2^2 - \omega^2)},$$

$$\alpha\delta k^2 = -\frac{al^2}{(\omega_*^2(1-\alpha_{g0}) - \omega^2)l^2 - a(N_2^2 - \omega^2)}$$

and

$$\beta^2 = K_1 l^2 + K_2,$$

where K_1 and K_2 are given by

$$K_1 = \frac{\omega^2 - N_1^2}{N_2^2 - \omega^2} + \frac{ag^2}{(N_2^2 - \omega^2)(\omega_*^2(1-\alpha_{g0}) - \omega^2)},$$

$$K_2 = \frac{a(N_1^2 - \omega^2)}{\omega_*^2(1-\alpha_{g0}) - \omega^2}.$$

Thus, for a given frequency ω, the dispersion relation (49) is an equation for l. In the limiting case of $h_b \to h$, (49) gives us the dispersion relation for a two-layer model, obtained earlier (Grimshaw and Khusnutdinova, 2004). In the limiting case of $h_b \to 0$ (or $h_b \to h$ and $\alpha\delta \to 0$), (49) gives us the dispersion relation for a two-layer pure fluid (no bubbles).

Next, we assume that $|\beta| h_b \ll 1$ (the upper bubbly layer is thin) and

$$vl \ll 1, \quad \text{where} \quad v = \frac{\omega(h-h_b)}{\sqrt{N_2^2 - \omega^2}} \tag{50}$$

(for $\omega < N_2$ this means that the vertical wavenumber l is small compared to v^{-1}). Note that if $h - h_b \to 0$, then $v^{-1} \to \infty$. Then (49) can be approximated by

$$b_1[\tan(H-h)l + (h-h_b)l] + h_b\left(b_2 l - \frac{b_3}{l}\right) = 0, \tag{51}$$

where the coefficients are given by

$$b_1 = \omega_*^2(1-\alpha_{g0}) - \omega^2 + agh_b,$$
$$b_2 = \omega_*^2(1-\alpha_{g0}) - \omega^2, \quad b_3 = a(N_2^2 - \omega^2).$$

For a given frequency ω and the given system parameters, the values of l determined by (51) are restricted to a discrete set of values (which depends on the frequency and system parameters).

For $h/H \ll 1$ we further simplify (51) and obtain the approximate solutions. Up to the first correction in h_b/H these solutions are given by

$$l_n \approx l_n^0 + \frac{h_b}{H}\tilde{l}_n, \quad n = 1, 2, \ldots, \tag{52}$$

where

$$\tilde{l}_n = \frac{a l_n^0}{b_1}\left(\frac{N_2^2 - \omega^2}{l_n^{02}} + gh_b\right), \quad l_n^0 = \frac{n\pi}{H},$$

together with the additional value l_0 given by

$$l_0^2 \approx \frac{b_3 h_b}{b_1 H - agh_b^2}$$

which can be interpreted as a bubble mode, as it does not exist if the upper layer has no bubbles. The modes with $n \geqslant 1$ can be interpreted as internal wave modes, which all exist when there are no bubbles. Note that these approximate solutions depend on the depth of the bubble layer and do not depend on the depth of the upper mixed layer. Therefore, the results obtained in our previous paper (Grimshaw and Khusnutdinova, 2004) remain valid. However, the region of the validity of these results is now defined more precisely, by imposing the additional condition (50). Note that the value of l_n, for a given frequency $\omega < N_2$ is increased due to the presence of the upper layer when bubbles are present (as then $\tilde{l}_n > 0$). It is interesting to note that if there are no bubbles in the upper layer, then $\tilde{l}_n$ is zero, and the correction to the leading order term $n\pi/H$ for l_n is then $O(h^3/H^3)$ ($l_n \approx n\pi/H + (n\pi)^3 h^3/3H^4$, so that the correction is still positive).

For each mode, the dispersion relation expressing ω as a function of k is obtained by combining expressions (41) and (52). It can then be solved for ω as a function of k. For the bubble mode ($l = l_0$) this gives the approximate relation

$$\omega^2 = \frac{[\omega_*^2(1-\alpha_{g0})H + agh_b(H - h_b)]k^2}{ah_b + Hk^2}. \tag{53}$$

Note that, as k^2 increases, so does ω^2, and hence there is a critical value k_N (obtained by setting $\omega^2 = N_2^2$ in (53)), such that for $k > k_N$, $l_0^2 < 0$ and the mode is evanescent in the lower layer, but for $k < k_N$, $l_0^2 > 0$ and the mode is then sinusoidal in the lower layer. In practice $k_N H \ll 1$ and so this latter possibility is confined to very long waves.

For internal waves in the lower layer, the dispersion relation is just (41) with $l \approx n\pi/H$ to leading order in h_b/H. Thus for these modes, $l^2 > 0$ and so $\omega^2 < N_2^2$. The small correction given by $\tilde{l}_n$ (52) has the effect of slightly increasing the wavenumber k, for a given ω, due to the presence of the bubbles in the upper layer, or equivalently of decreasing the frequency ω for a given wavenumber k. Further, this effect is $O(h_b/H)$ and for high void fractions and small ratios h/H it may be comparable with the corresponding effect in the absence of any bubbles, when it is $O(h^3/H^3)$.For example, if $h_b = 10\,\mathrm{m}$, $h = 50\,\mathrm{m}$, $H = 4000\,\mathrm{m}$, $R_0 = 100\,\mu\mathrm{m}$, $\omega_* = 2.5\times10^5\,\mathrm{s}^{-1}$, $N_2 = 10^{-3}\,\mathrm{s}^{-1}$, $\omega = 0.5N_2$, we obtain $(l_1 - l_1^0)/l_1^0 \approx 1.2\times10^{-3}\,\alpha_{g0}$, and this can be comparable with the corresponding correction in the absence of any bubbles, when it is $O(10^{-6})$ (the corresponding value in the framework of the two-layer model with $h = h_b = 10\,\mathrm{m}$ was $O(10^{-8})$). On the other hand, in shallow water, for $H = 100\,\mathrm{m}$ and $h = h_b = 10\,\mathrm{m}$ say, $(l_1 - l_1^0)/l_1^0 \approx 5\times10^{-2}\,\alpha_{g0}$. This is smaller than the corresponding correction in the absence of any bubbles, which is $O(10^{-3})$, but can also be significant if the void fraction is high enough.

Next we consider the case when N_2 is zero, so that in effect the stratification is confined to the upper bubble layer, and is due to the depth-dependent distribution of bubbles only. We then have $l = ik$. We also now relax the hypothesis that $|\beta|h_b \ll 1$. To allow for the waves to be sinusoidal in the upper bubble layer we now set $\beta = i\tilde{\beta}$. The dispersion relation (49) then takes

the form

$$(1-\alpha\delta\omega^2)\tilde{\beta}\cot\tilde{\beta}h_b + k\coth(H-h_b)k = \alpha\delta gk^2, \tag{54}$$

where, from (45),

$$k^2 = K_1^{-1}\tilde{\beta}^2 + \frac{K_2}{K_1}. \tag{55}$$

Therefore, the dispersion relation (54) has a structure

$$\tilde{\beta}\cot\tilde{\beta}h_b = F(\tilde{\beta},\omega),$$

where, for each fixed ω and for all $k^2>0$, F is a continuous function of $\tilde{\beta}^2$. Since $\cot\tilde{\beta}h_b$ will pass from $+\infty$ to $-\infty$ an infinite number of times as $\tilde{\beta}h_b$ varies, we can conclude that there will be an infinite number of discrete modes in the upper (bubbly) layer. For example, in the limiting case of $H\to h\to h_b$ (i.e. there is only the bubble layer), (54) gives us $\sin\tilde{\beta}h_b=0$, so that $\tilde{\beta}_n = n\pi/h_b$, $n=1,2,3\ldots$. We now notice, that since $N_2=0$, then (55) coincides with the dispersion relation (29) now solved for k^2 as a function of the frequency ω and the vertical wavenumber $\tilde{\beta}$. Therefore, the eigenfrequencies are given by the dispersion relation found in Section 3, where the vertical wavenumber $\tilde{\beta}$ is quantized to the restricted set of values defined by (54). For example, in the limiting case of $H\to h\to h_b$, for each value of $\tilde{\beta}$ there are two branches of the dispersion relation, the "internal" wave and the "bubble" wave (see Section 3). For the "internal" wave, the frequency lies below a value slightly lower than N_1; for the "bubble" wave, the frequency is close to N_1 for very long waves, but typically, is much higher (see Section 3 for details). Note that N_1 is the effective buoyancy frequency in the upper (bubbly) layer, and since we assume here that there is no background stratification, the effective buoyancy frequency is given by

$$N_1^2 \approx g\alpha_{g0}(\ln n_0)_z.$$

For example, for $\alpha_{g0}=5\times10^{-6}$, $N_1\approx 7\times10^{-3}\,\mathrm{s}^{-1}$. Thus, the presence of a depth-dependent distribution of bubbles alone leads to the possibility of existence of two classes of waves (bubble waves and internal waves) in the upper (bubbly) layer.

5. Conclusion

In this paper we have described a study of the influence of an upper bubbly layer on the properties of the oceanic waveguide. We have considered a simplified configuration, in which the horizontal variability of bubble distributions in a real oceanic environment has been ignored in favour of a study of the consequences of the depth-dependent distribution of bubbles. Further, in this simplified model, the depth-dependence of the actual bubble layer was taken into account only through the effective buoyancy frequency, while the remaining variable coefficients in the equations of motion were replaced by their average values over the depth of the bubble layer. A three-layer model has been used to account for the mismatch between the depth of the bubble layer and the depth of the upper mixed layer. As follows from our analysis, the presence of the vertically variable distribution of bubbles introduces an effective stratification and essentially changes the value of the buoyancy frequency in the upper part of the ocean, which should be replaced by an effective value. Further, the presence of the upper bubbly layer causes a major change to the dispersion relation, in that a "bubble" mode is introduced along with the internal waves, which typically has a higher frequency than the buoyancy frequency (except possibly for very long waves). Also, we have shown that the presence of a depth-dependent distribution of bubbles alone leads to the possibility of existence of two classes of waves (bubble waves and internal waves) in the upper (bubbly) layer. There is usually a spectral gap between the "bubble" waves and the internal waves in the bubbly layer, with the effective buoyancy frequency being a lower bound for the "bubble" wave, and the upper bound for the internal wave being a bit less than the effective buoyancy frequency. All this points to the essential influence of the bubble distributions in the upper ocean on the properties of the internal wave field and invites further detailed study.

Acknowledgements

We are thankful to S.L. Gavrilyuk, R.I. Nigmatulin and L.A. Ostrovsky for useful and

stimulating discussions. We appreciate support of the Department of Mathematical Sciences of Loughborough University.

References

d'Agostino, L., Brennen, C.E., 1989. Linearized dynamics of spherical bubble clouds. Journal of Fluid Mechanics 199, 155–176.

Akhatov, I.Sh., 1989. The shear instability in stratified viscoelastic and liquid–gas media. Journal of Applied Mathematics and Mechanics 53 (4), 630–635.

Brennen, C.E., 1995. Cavitation and Bubble Dynamics. Oxford University Press, Oxford, 282pp.

Brennen, C.E., 1998. Cloud cavitation: observations, calculations and shock waves. Multiphase Science and Technology 10, 303–321.

Buckingham, M.J., 1997. Sound speed and void fraction profiles in the sea surface bubble layer. Applied Acoustics 51 (3), 225–250.

Bullock, G.N., Crawford, A.R., Hewson, P.J., Walkden, M.J.A., Bird, P.A.D., 2001. The influence of air and scale on wave impact pressures. Coastal Engineering 42, 291–312.

Crawford, G.B., Farmer, D.M., 1987. On the spatial distribution of ocean bubbles. Journal of Geophysical Research 92, 8231–8243.

Dean, G.B., Stokes, M.D., 1999. Air entrainment processes and bubble size distributions in the surf zone. Journal of Physical Oceanography 29, 1393–1403.

Druzhinin, O.A., Ostrovsky, L.A., Prosperetti, A., 1996. Low-frequency acoustic wave generation in a resonant bubble-layer. Journal of the Acoustical Society of America 100, 3570.

Farmer, D.M., Lemon, D.D., 1984. The influence of bubbles on ambient noise in the ocean at high wind speeds. Journal of Physical Oceanography 14, 1762–1778.

Farmer, D.M., Vagle, S., 1989. Waveguide propagation of ambient sound in the ocean-surface bubble layer. Journal of the Acoustical Society of America 86, 1897–1908.

Farmer, D.M., Vagle, S., Booth, A.D., 1997. A free-flooding acoustical resonator for measurement of bubble size distributions. Journal of Atmospheric and Oceanic Technology 15, 1132–1146.

Farmer, D.M., Vagle, S., Li, M., 1999. Wave breaking, turbulence and bubble distributions in the ocean surface layer. In: Banner, M.L. (Ed.), Proceedings, The Wind-driven Air–sea Interface. The University of New South Wales, Sydney, Australia, pp. 187–192.

Garrett, C., Li, M., Farmer, D.M., 2000. The connection between bubble size and energy dissipation rates in the upper ocean. Journal of Physical Oceanography 30, 2163–2171.

Gill, A.E., 1982. Atmosphere—Ocean Dynamics. Academic Press, New York, 662pp.

Grimshaw, R.H.J., Khusnutdinova, K.R., 2004. The effect of bubbles on internal waves. Journal of Physical Oceanography 34, 477–489.

Iordansky, S.V., 1960. Equations of motion of liquid containing gas bubbles. Zhurnal Prikladnoj Mekhaniki Tekhnicheskoi Fiziki 3, 102–110 (in Russian).

Johnson, B.D., Cooke, R.C., 1979. Bubble populations and spectra in coastal waters: a photographic approach. Journal of Geophysical Research 84, 3761–3766.

Karpov, S., Prosperetti, A., Ostrovsky, L., 2003. Nonlinear wave interactions in bubble layers. Journal of the Acoustical Society of America 113 (3), 1304–1316.

Kogarko, B.S., 1961. On the model of cavitating liquid. Doklady AN SSSR 137, 1331–1333 (in Russian).

Kolovaev, P.A., 1976. Investigation of the concentration and statistical size distribution of wind-produced bubbles in the near surface ocean layer. Oceanology 15, 659–661.

Kumar, S., Brennen, C.E., 1993. Some nonlinear interactive effects in bubbly clouds. Journal of Fluid Mechanics 253, 565 591.

Kutateladze, S.S., Nakoryakov, V.E., 1984. Heat and mass-transfer and waves in gas–liquid systems. Nauka, 301pp (in Russian).

Medwin, H., 1977. Acoustic fluctuations due to microbubbles in the near-surface ocean. Journal of the Acoustical Society of America 56, 1100–1104.

Medwin, H., 1977. In situ acoustic measurements of micro-bubbles at sea. Journal of Geophysical Research 82, 971–976.

Melville, W.K., 1996. The role of surface-wave breaking in air–sea interaction. Annual Review of Fluid Mechanics 28, 279–321.

Minnaert, M., 1933. On musical air bubbles and the sounds of running water. Philosophical Magazine 16, 235–248.

Nakoryakov, V.E., Pokusaev, V.G., Shreiber, I.R., 1983. Propagation of Waves in Gas– and Vapour–liquid Media. Institute of Thermorhysics, Novosibirsk, 280pp (in Russian).

Naugolnykh, K., Ostrovsky, L., 1998. Nonlinear Wave Processes in Acoustics. Cambridge University Press, Cambridge, 298pp.

Nigmatulin, R.I., 1991. Dynamics of Multiphase Media, vol. 1 and 2. Hemisphere, Washington, DC, 507 and 371pp.

Ostrovsky, L.A., Sutin, A.M., Soustova, I.A., Matveyev, A.I., Potapov, A.I., 1998. Nonlinear, low-frequency sound generation in a bubble layer: theory and laboratory experiment. Journal of the Acoustical Society of America 104, 722–726.

Plesset, M.S., Prosperetti, A., 1977. Bubble dynamics and cavitation. Annual Review Fluid Mechanics 9, 145–185.

Prosperetti, A., 1994. Bubble dynamics: some things we did not know 10 years ago. In: Blake, J., Boulton-Stone, J., Tomas, N. (Eds.), Bubble Dynamics and Interface Phenomena. Kluwer, Dordrecht, pp. 3–16.

Rayleigh, Lord, 1917. The pressure developed in a liquid on the collapse of a spherical cavity. Philosophical Magazine 34, 94.

Reisman, G.E., Wang, Y.C., Brennen, C.E., 1998. Observation of shock waves in cloud cavitation. Journal of Fluid Mechanics 355, 255 283.

Scott, J.C., 1975. The role of salt in whitecap persistence. Deep-Sea Research 22, 653–657.

Terrill, E.J., Melville, W.K., 1999. Field measurements of bubble size distributions in the upper mixed layer. In: Banner, M.L. (Ed.), Proceedings, The Wind-driven Air–Sea Interface. The University of New South Wales, Sydney, Australia, pp. 247–255.

Thorpe, S.A., 1982. On the clouds of bubbles formed by breaking wind-waves in deep water, and their role in air–sea gas transfer. Philosophical Transactions of the Royal Society of London A 304, 155–210.

Thorpe, S.A., 1984a. On the determination of K_v in the near-surface ocean from acoustic measurments of bubbles. Journal of Physical Oceanography 14, 855–863.

Thorpe, S.A., 1984b. A model of the turbulent diffusion of bubbles below the sea surface. Journal of Physical Oceanography 14, 841–853.

Thorpe, S.A., 1995. Dynamical processes of transfer at the sea surface. Progress in Oceanography 35, 315–352.

Thorpe, S.A., Hall, A.J., 1983. The characteristics of breaking waves, bubble clouds, and near-surface currents observed using sidescan sonar. Continental Shelf Research 1 (4), 353 384.

Thorpe, S.A., Belloul, M.B., Hall, A.J., 1987. Internal waves and whitecaps. Nature 330, 740–742.

Urick, R.J., 1983. Principles of Underwater Sound. McGraw-Hill, New York, 384pp.

van Wijngaarden, L., 1968. On the equations of motion for mixtures of liquid and gas bubbles. Journal of Fluid Mechanics 33, 465–474.

Woolf, D.K., Thorpe, S.A., 1991. Bubbles and the air–sea exchange of gases in near saturation conditions. Journal of Marine Research 49, 435–466.

Zedel, L., Farmer, D., 1991. Organized structures in subsurface bubble clouds: Langmuir circulation in the open ocean. Journal of Geophysical Research 96 (C5), 8889–8900.

ELSEVIER

Deep-Sea Research II 51 (2004) 2919–2941

DEEP-SEA RESEARCH
PART II

www.elsevier.com/locate/dsr2

Intermittency of vertical density gradients at finescale and link with mixing processes

Pascale Bouruet-Aubertot[a,*], J. Sommeria[b], B. Le Cann[c], C.R. Koudella[d]

[a]*LODYC/IPSL, Université Pierre et Marie Curie/UMR 7617, Paris, France*
[b]*LEGI, Université Joseph Fourier, Grenoble, France*
[c]*LPO, Université de Bretagne occidentale, Brest, France*
[d]*DEAS, Harvard University, Cambridge, USA*

Abstract

Nonlinear interactions among internal gravity waves lead to intermittent breaking events with density overturning and vertical turbulent mixing. We analyze the statistics of density fluctuations generated in this process, using direct numerical simulations. We then check the relevance of these results for in situ density profiles measured during the ARCANE experiment on the continental slope west of the Iberian peninsula, a region with strong internal tides. We focus the analysis in the depth range [400–1000 m] for which the temperature is uniform, and the density stratification is due mainly to salinity.

We verify in the numerical simulations that the histograms of vertical density gradients become strongly skewed at small vertical scales. This intermittency of density fluctuations can be attributed to the formation of sheets with strong stable density gradients. These sheets result from convergent motion produced by surrounding breaking events. We check that the vertical density gradients at finescale can be represented by a model of layers and sheets distributed with a Poisson statistics, with a mean number of sheets per unit length, μ. We next show that the mean sheet spacing μ^{-1} is proportional to the Ozmidov scale (proportional to the square root of the energy dissipation). Using the classical link between energy dissipation and vertical eddy diffusivity K_d, we deduce a link between the statistics of vertical density gradients and the vertical mixing, namely K_d is proportional to $N\mu^{-2}$.

For in situ measurements, we find a similar statistics of vertical density gradients, confirming earlier results obtained by Hayes et al. [J. Geophys. Res. 80 (3) 314] and Pinkel and Anderson [J. Phys. Oceanogr. 22 (1992) 773] in the open ocean thermocline. This allows us to estimate a rate of total energy dissipation 2.4×10^{-8} W/kg and a vertical eddy diffusivity 5×10^{-4} m^2 s^{-1}. Comparisons with other estimates of energy dissipation or vertical mixing would be required to validate this method.
© 2004 Elsevier Ltd. All rights reserved.

*Corresponding author.
E-mail address: pba@lodyc.jussieu.fr (P. Bouruet-Aubertot).

0967-0645/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dsr2.2004.09.007

1. Introduction

Wavebreaking corresponds to the onset of turbulence, with irreversible energy dissipation, from an initially smooth wave field. It controls mixing in stably stratified fluids, as other sources of turbulence tend to be suppressed. In many parts of the deep ocean, a quasi-permanent level of random internal wave activity is maintained by various sources. Internal waves are in particular produced from baroclinic tides, themselves generated by the interaction of the main (barotropic) tides with the bottom topography. They are also produced by atmospheric forcing, especially during storms or cooling events. Steepening by wave–wave interactions and refraction by density fluctuations produces a cascade of wave energy to smaller scales with stronger nonlinearity until breaking occurs. Turbulence is thus sporadically produced, leading to vertical mixing of heat and solutant.

Note that internal waves are also important in the atmosphere. They are emitted by the effect of topography or storms and are amplified by air rarefaction as they propagate upward until breaking occurs. This has a significant influence on the vertical wind profile, as momentum from the lower atmosphere is transferred to the altitude of breaking. The resulting regimes of wave induced turbulence have common properties in the atmosphere and oceans.

These properties have been reproduced to some extent in laboratory experiments by Bénielli and Sommeria (1998) and in numerical simulations by Bouruet-Aubertot et al. (1996) and Carnevale et al. (2001). In these studies the internal waves are produced in a uniformly stratified fluid with no Coriolis effects. Furthermore, turbulence is induced by the breaking of a simple wave which is initially introduced, or permanently forced at large scale. These studies had primarily focused on spectral analysis and global energy dissipation and mixing effects. The goal of the present paper is twofold. The first one is to characterize the intermittency of the vertical density structure in such well controlled configurations and to relate these processes with mixing. The second is to apply these results to oceanic data and to provide an indirect estimate of mixing.

The wave induced turbulence is characterized by an energy spectrum scaling as k_z^{-3} along the vertical as observed both in the atmosphere (e.g. Alisse and Sidi, 2000), and in the ocean (e.g. Gargett et al., 1981). It connects the weakly nonlinear waves at the larger scales to the fully turbulent but sporadic micro-scales. This spectral domain is, therefore, important in determining the mixing properties of the internal wave field. It is called the buoyancy sub-range in the atmosphere (ranging from 100 m to a few km in vertical scales). In the oceans it corresponds to the so-called fine structure, or compliant waves (for vertical scales ranging from 1 to 10 m). Such a k_z^{-3} spectral range has been reproduced in the laboratory experiments of Bénielli and Sommeria (1998) and in the numerical simulations of (Bouruet-Aubertot et al., 1996; Carnevale et al., 2001).

Intermittency is known to be important in this wave induced turbulence. Sporadic occurrence of well mixed zones or strongly stratified layers have been documented. Baker and Gibson (1987) pointed out the requirement to take into account such intermittency when analyzing oceanic data of energy dissipation and mixing. However, the question of lognormality of dissipation rates in the ocean should be examined with care as underlined by Yamazaki and Lueck (1990) who proposed a robust method of computation of these quantities.

For weakly interacting internal wave fields we expect Gaussian statistics for both velocity and buoyancy. This results from general arguments of random phase (e.g. Olbers, 1976). The statistics then becomes more and more intermittent at smaller and smaller wavelengths, until the scale of the sporadic breaking events. This has been observed in several oceanic campaigns (Hayes et al., 1975; Desaubies and Gregg, 1981; Pinkel and Anderson, 1992), and recently confirmed in direct numerical simulations (Bouruet-Aubertot et al., 2004): while Gaussian distributions are obtained for buoyancy and velocity fluctuations at scales corresponding to resonantly interacting waves, skewed distributions are obtained for density fluctuations in the k_z^{-3} range, becoming more and more skewed at smaller scales, corresponding to the rare occurrence of stratification steps. The energy

dissipation is also intermittent: Gregg et al. (1986, 1993) found a lognormal distribution for the dissipation rate in the ocean (however perturbed by instrumental noise and undersampling).

Theoretical models have been developed in order to explain this strong skewness of the density structure. However, the link between energy dissipation (related to vertical mixing) and density statistics remains unclear.

Desaubies and Gregg (1981) have shown that the background temperature gradient is locally increased or decreased in an asymmetric way by the transport of a Gaussian wave field without energy dissipation. Their model provides an adequate description of oceanic observations for vertical scales larger than 2 m. However it fails to describe smaller scales, at which breaking events and mixing occur.

An opposite model has been proposed by Hayes et al. (1975), which assumes that turbulent mixing cuts random steps in the initial uniform vertical stratification: each uniform layer is topped by a discontinuity sheet to keep up with the background gradient. The step edges are assumed randomly located, which results in Poisson statistics. Hayes et al. (1975) found a reasonable agreement with their oceanic data in the relevant range of vertical scales (1–10) m. Evidence for such stepwise structures has also been given in the atmosphere (Dalaudier et al., 1994).

Interestingly this vertical density structure has been retrieved from numerical simulations of internal gravity wavebreaking (Bouruet-Aubertot et al., 2004), opening the opportunity to test accurately the range of validity of these two models. It was found that the strain model is adequate for scales larger than 30 times the Ozmidov length scale, while the Poisson statistics provides a fairly good description for smaller scales. This should be viewed as a convenient practical description, although the underlying step and layer model is not realistic. Pinkel and Anderson (1992) have derived for instance the same Poisson statistics from an alternative physical model (a pile of layers with random density gradients). Taking into account strain variability as a function of vertical separation they later derived a model for the pdf of the Richardson number (Pinkel and Anderson, 1997a) and then inferred an eddy diffusivity (Pinkel and Anderson, 1997b).

The key question addressed in this paper is to relate this statistics of density to mixing processes. To this aim we perform numerical computations that start with a simple initial condition (a large-scale internal gravity wave). In this way, turbulence is produced by an intrinsic sequence of instabilities without external noise. A preliminary analysis of in situ oceanic measurements checking the relevance of the numerical results is also presented.

The paper is organized as follows: in Section 2 we describe the numerical model and introduce the physical parameters. In Section 3 we introduce the main characteristics of the wave induced turbulence and the relevant turbulent length scales. We next give evidence of the intermittency of the vertical density structure at small scales and characterize the most intense events that take the form of strongly stratified sheets in Section 4. In Section 5 we show how mixing can be inferred from the statistics of these stratified sheets. We finally analyze salinity and temperature profiles taken during the ARCANE experiments in Section 6.

2. Numerical model

We solve the Navier–Stokes equations in the Boussinesq approximation, either in the vertical plane (2D computations), or in three dimensions (3D computations). The domain is accordingly a vertical square or a cube. We use Cartesian coordinates (x, z) or (x, y, z), with the vertical coordinate z directed upwards. The basic density profile, $\bar{\rho}$, is linearly decreasing with the vertical coordinate z and the buoyancy frequency

$$N^2 = -\frac{g}{\rho_0}\frac{\mathrm{d}\bar{\rho}}{\mathrm{d}z},$$

is thus constant. In the Boussinesq approximation, the density variations are assumed to be small with respect to the mean density ρ_0.

In two dimensions, the condition of incompressibility is satisfied by introducing a stream

function Ψ, related to the two velocity components u and w by $u = \partial_z \Psi$, $w = -\partial_x \Psi$. We solve the Boussinesq equations in terms of the vorticity $-\Delta\Psi$ and the buoyancy fluctuations ρ^*: the actual density is therefore $\rho = \rho_0 + (\rho_0/g)\rho^*$.

$$\frac{\partial \Delta\Psi}{\partial t} + J(\Delta\Psi, \Psi) = \frac{\partial \rho^*}{\partial x} + \nu\nabla^2\Psi,$$
$$\frac{\partial \rho^*}{\partial t} + J(\rho^*, \Psi) = -N^2\frac{\partial \Psi}{\partial x} + \kappa\nabla^2\rho^*,$$

where the non linear advective terms are written by means of the Jacobian J and υ is the viscosity. The density is assumed to be a linear function of the transported quantity, temperature or salinity, with diffusivity κ, the Prandtl number $Pr = (\upsilon/\kappa)$ is here taken equal to unity. We study standing waves in a square of side π and we use periodic boundary conditions with symmetry (e.g. Canuto et al., 1988). The normal velocities at the boundaries are set to zero as well as density fluctuations at the horizontal boundaries, $z = 0$ and π.

In three dimensions, we solve the equations of motion, with the Boussinesq approximation:

$$\frac{\partial \vec{v}}{\partial t} + \vec{v}.\nabla\vec{v} = -\frac{\vec{\nabla}p'}{\rho_0} + \frac{\rho'}{\rho_0}\vec{g} + \nu\nabla^2\vec{v},$$
$$\nabla.\vec{v} = 0,$$
$$\frac{\partial \rho'}{\partial t} + \vec{v}.\nabla\rho' = \frac{\rho_0}{g}N^2 w + \kappa\nabla^2\rho',$$

where p' is the pressure fluctuation divided by ρ_0 and $\vec{i}_z$ is the unit vector along the vertical direction oriented upward. As in the 2D case, we only consider standing waves with different amplitudes.

The parameters of the runs are given in Table 1. We analyzed mainly two cases. The first one is that of two-dimensional simulations with resolution 512^2 and the second one is that of three-dimensional simulations, with lower resolution 128^3, of the same wave with a perturbation superimposed.

The results of our calculations can be easily transposed to oceanic and atmospheric flows having the same Froude number and Reynolds number as those of our calculations. We look here at the evolution of a large-scale primary wave, of wavelength equal to $\lambda = 2\pi$ in the x and y directions, and which is evolving in a background linear stratification, of buoyancy frequency N. The numerical domain has an area equal to $\lambda^2/4$ in the case of standing waves. Comparisons with oceanic or atmospheric flows can be made through a similarity in Froude number Fr, and Reynolds number Re. Fr and Re are defined from the primary wave parameters and background stratification: $k_z U/N$ and $Re = (UL/\upsilon)$, where U is the amplitude of the horizontal velocity of the primary wave. We shall represent our results in terms of dimensionless quantities to facilitate comparisons. As an example of application of our simulations to an oceanic situation, let us consider an internal wave propagating in the thermocline, with $Fr = 0.256$ and $Re = 1.1 \times 10^4$ (as for run 2 in Table 1). These parameters may characterize a wave with wavelength $\lambda = 20\,\mathrm{m}$ and frequency $N/\sqrt{2}$ evolving in a background stratification $N = 2.5 \times 10^{-3}\,\mathrm{rad/s}$. Note that the Reynolds number of this oceanic wave is slightly higher than that of the numerical simulation ($Re = 3 \times 10^4$ compared to $Re = 1.1 \times 10^4$).

However, this difference in Reynolds numbers has no effect on the dynamics which is controlled by nonlinear wave–wave interactions and is therefore prescribed by the value of the Froude number.

Table 1
Physical and numerical parameters of the simulations: $Fr = k_z U/N = k_z^2 A/N$ is the Froude number of the primary wave, it can be alternatively defined from the amplitude of the horizontal velocity U or the amplitude of the stream function and the vertical wavenumber of the primary wave, Fr equals here the amplitude A since $k_z = 1$ and $N = 1$, ν is the viscosity introduced in the model and Re is the Reynolds number based upon the horizontal velocity scale and the size of the domain, $Re = UL/\nu$

	Resolution	$Fr = k_z U/N$	ν	Re
Run1	129^3	0.256	3×10^{-5}	8.5×10^3
Run2	513^2	0.256	2.4×10^{-5}	1.1×10^4
Run3	129^2	0.256	5×10^{-5}	5.1×10^3
Run4	513^2	0.384	4×10^{-5}	9.6×10^3
Run5	513^2	0.448	4×10^{-5}	1.1×10^3
Run6	129^3	0.512	4×10^{-5}	1.3×10^3
Run7	513^2	0.512	4×10^{-5}	1.3×10^3

3. Wave evolution toward breaking and induced turbulence

The initial condition is chosen to be a standing wave, solution of the linear Boussinesq equations

$$\Psi(x,z) = A\sin(k_x x)\sin(k_z z) = \frac{FrN}{k_z^2}\sin(k_x x)\sin(k_z z), \quad \rho^* = 0.$$

It oscillates at the frequency ω given by the linear dispersion relationship $\omega = N\cos(\theta)$, where θ is the angle between the wave-vector and the horizontal plane. We choose $k_x = k_z = 1$, so that $\omega = N/\sqrt{2}$. The nonlinearity of the primary wave is characterized by the Froude number $Fr = \frac{k_z U}{N}$, where $U = k_z A$ is the amplitude of the horizontal velocity. Alternatively we could have defined a Froude number that corresponds to the maximum steepness of the isopycnals: $\frac{2k_x kA}{\pi N}$. In the three-dimensional simulations, an additional wave perturbation with an amplitude equal to $\frac{1}{20}A$ and of wave vector $\vec{k} = (0,1,1)$ is introduced in order to favour the onset of fully three-dimensional wavebreaking.

Since a standing wave is only a solution of the linearized equations, the actual solution progressively departs from the linear approximation, involving weakly nonlinear interactions, followed by a breaking process (see Bouruet-Aubertot et al., 1995). The turbulent regime of interest is reached after wave breaking, (as discussed by Bouruet-Aubertot et al., 1996), when energy is decaying by energy transfer towards small scales (Fig. 1). During this turbulent regime, local density inversions (overturning regions) coexist with strongly stratified regions.

Because of the anisotropy due to the stratification, we distinguish between spectra along the vertical and horizontal directions. Both display an

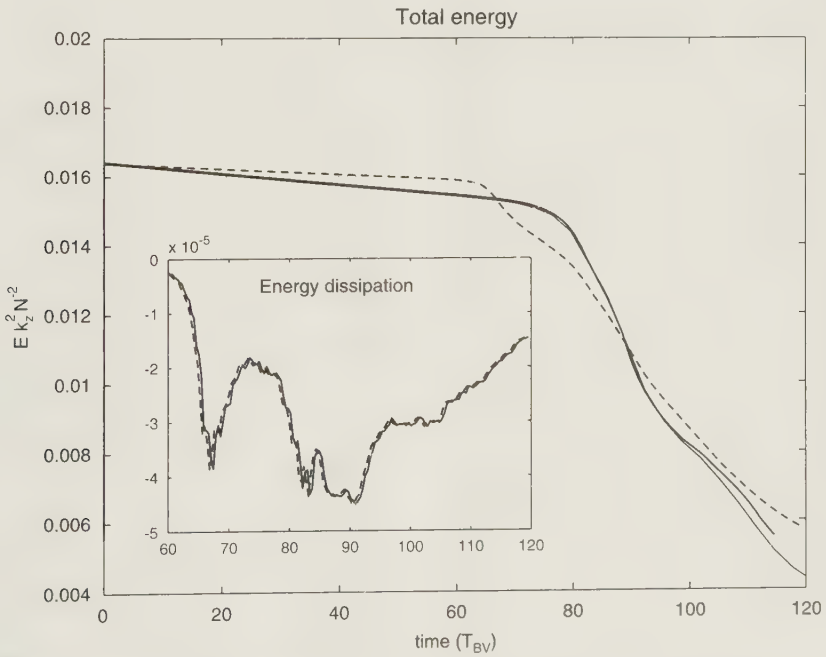


Fig. 1. Time evolution of the total energy (kinetic energy + available potential energy), $E = 0.5[|\vec{v}|^2 + g^2/(\rho_0 N)^2\rho'^2]$, for an initial wave with a Froude number $Fr = 0.256$ (E is represented in a dimensionless form: $Ek_z^2N^{-2}$). The three-dimensional calculation is represented with a thick bold line (run1) while the two-dimensional calculations are represented with a thick dashed line for the 512^2 resolution (run2) and with a thin line for the 128^2 resolution (run3). The analysis is focused on the time interval $[83T_{BV}, 92T_{BV}]$ corresponding to a total energy within $[60\%E(t=0), 80\%E(t=0)]$. Time evolution of the energy time derivative is represented in a sub-figure (black line) as well as the dissipation rate computed directly (dashed line).

inertial range with a spectral slope close to -3. It was shown by Bouruet-Aubertot et al. (1996) that the spectral energy at a wavenumber k_z (calculated along the vertical direction) saturates at a level $\alpha N^2 k_z^{-3}$ and does not depend on the initial condition. A relevant scale for this buoyancy subrange is the Ozmidov's scale λ_0, defined from the rate of energy dissipation ε by

$$\lambda_0 = \left(\frac{\varepsilon}{N^3}\right)^{1/2}.$$

We calculate ε as the total energy drop ΔE (per unit of volume) during the time of analysis Δt (namely $E \approx [60\%E(t=0), 80\%E(t=0)]$), $\varepsilon = -(\Delta E/\Delta t)$.

An example of the energy decrease is shown in Fig. 1 for the smallest wave Froude number ($Fr = 0.256$). Three different simulations are shown: a low resolution three-dimensional simulation, the corresponding two-dimensional simulation with the same resolution and a high resolution two-dimensional simulation. The first two curves are very close which demonstrates that overall nonlinear energy transfers have the same global effect on the total energy though details in kinetic and potential energy transfers might differ between the two-dimensional and the three-dimensional case. The two-dimensional higher resolution case matches the two other cases closely as well, the main difference arising from the viscosity introduced in the model (see Table 1): the onset of wavebreaking occurs earlier in the latter case (2D high resolution) because of the smaller viscosity of the simulation. Later on, during the time interval of interest, the total energy decreases in a very similar way in all three cases. It results from non linear energy transfers that lead to wavebreaking. Because these transfers are set at large scales by the instability of the primary wave, the resulting energy decrease is not strongly modified whether two-dimensional or three-dimensional simulations are considered. In contrast fully 3D dynamics is obtained when the large scale primary wave is statically unstable as shown by Koudella (1999). This finding is well predicted by Floquet stability analysis (Klostermeyer, 1991; Lombard and Riley, 1996) which shows that 2D perturbations are more unstable than 3D perturbations so far as the primary wave is statically stable.

The rate of energy decay ε is plotted in Fig. 2 versus the initial wave Froude number Fr. The Ozmidov scale $\lambda_0 = (\varepsilon/N^3)^{1/2}$ is represented. Its variation is consistent with a law in $Fr^{3/2}$ (ε in Fr^3). We notice again that the two-dimensional computations (empty circles) are in good agreement with the three-dimensional ones (filled circles).

Therefore, we analyze both types of simulations, the three-dimensional case allows us to investigate the realistic structure of events observed during wavebreaking, while the high resolution two-dimensional simulations enable us to investigate intermittency at small vertical scales and provide a link with mixing processes.

4. Evidence for strongly stratified sheets

4.1. Histograms of density differences at different vertical scales

We compute the histograms (probability distribution functions, pdf) of density differences, $(\rho(z+\Delta z) - \rho(z))/\Delta z$ for different vertical scales, Δz. Thus, neutral stratification corresponds to the zero abscissa in the figures and the mean stratification corresponds to -1. These pdfs are displayed in Fig. 3 for different Δz and for two Froude numbers of the wave, $Fr = 0.256$ and 0.512. Both three- and two-dimensional simulations are shown. Note that a comparison between these cases is made for similar vertical separations in terms of units of the Ozmidov length scale. This is consistent with the fact that the shape of the pdfs evolves as a function of the vertical separation as a different dynamical regime is concerned (from weakly nonlinear waves at large scales to breaking waves at smaller scales).

We notice the obvious asymmetry (strong skewness) of the pdf at small scales (Figs. 3A and C), which reflects the difference between stably stratified regions (left side of the curves) and overturned regions (right side). On the left side, long tails are obtained at small scales, corresponding to highly stratified sheets: a plateau is almost reached (Figs. 3A and B), with local

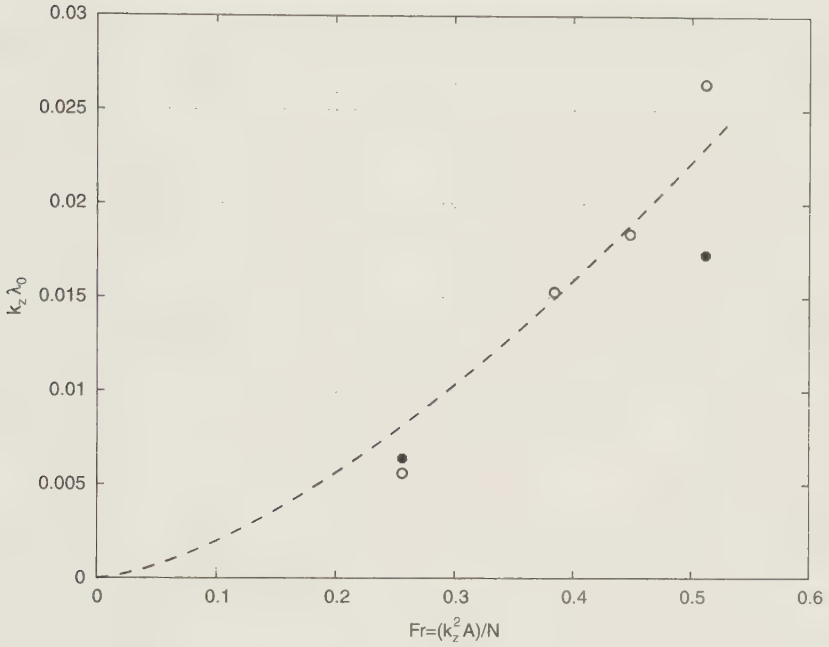


Fig. 2. Ozmidov length scale $\lambda_0 = (\varepsilon/N^3)^{0.5}$ as a function of wave Froude number Fr, 2D simulations are represented with empty circles and 3D simulations with filled circles. A $Fr^{3/2}$ fit is represented with a dotted line.

stratification reaching up to thirty times the initial background stratification (Fig. 3C). This asymmetry decreases as the vertical scale increases (Figs. 3B and D). Very similar results are obtained for the three-dimensional simulations especially when the primary wave is of small amplitude (Figs. 3A and B). This results from the fact that the dynamics remains almost two-dimensional even during wavebreaking (e.g. Fig. 3B and Bouruet-Aubertot et al., 2001). When the wave is of higher amplitude slight differences can be noticed (Figs. 3C and D) especially for strongly stratified regions. These differences can be attributed to convective instabilities that occur in the transverse plane in the three-dimensional simulations, as will be discussed in the next section.

In order to measure the probability of occurrence of overturning events, we have computed the relative volume of unstable (overturned) regions compared to the total volume of fluid. This quantity is displayed as a function of the vertical scale Δz in Fig. 4. The plot underlines the fact that breaking events are localized in space (small vertical extent) and rare (a maximum of 18% is reached for the highest wave amplitude). We expect that the strong stable density gradients are produced at the edge of regions mixed by local overturning. We check this assumption in the next section by a direct examination of the most intense events.

4.2. Dynamical processes associated with strongly stratified regions

We have selected the most intense events characterized by extreme stable density gradients (at scale $\Delta z \leqslant 10\lambda_0$). They appear as thin horizontal sheets surrounded by mixed convective zones. An example is visualized in Fig. 5 by means of maps of density and vertical density gradient. In the plane (x, z) of the primary wave the most strongly stratified region is located around $x \approx [2.5, 3.14]$ and $z \approx 0.5$ (Figs. 5A and C), with a vertical density gradient up to 15 times the background value. We notice local density overturning on both sides of this stratified sheet (visualized by dark

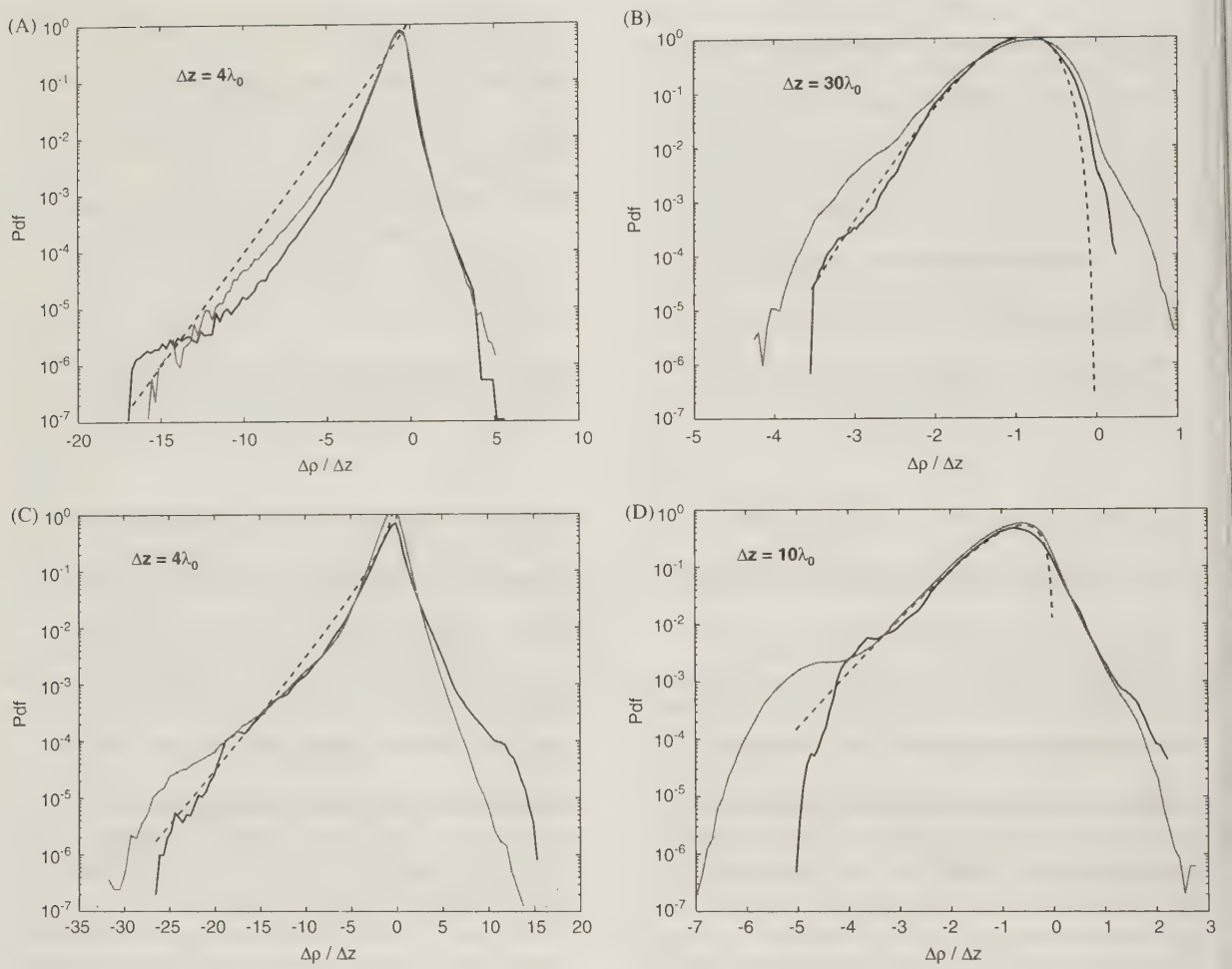


Fig. 3. Probability density function of the density gradient at different vertical scales pdf($[\rho(x, z + \Delta z) - \rho(x, z)]$)/$\Delta z$: (A) $\Delta z = 4\lambda_0$ (i.e. $\Delta z = 4\,\mathrm{d}z$), $Fr = 0.256$ (runs 1 and 2), (B) $\Delta z = 30\lambda_0$ (i.e. $\Delta z = 32\,\mathrm{d}z$), $Fr = 0.256$ (runs 1 and 2), (C) $\Delta z = \lambda_0$ (i.e. $\Delta z = 4\,\mathrm{d}z$ for the 3D simulation, and $\Delta z = 8\,\mathrm{d}z$ for the 2D simulation), $Fr = 0.512$ (runs 6 and 7), (D) $\Delta z = 10\lambda_0$, (i.e. $\Delta z = 32\,\mathrm{d}z$ for the 3D simulation, and $\Delta z = 48\,\mathrm{d}z$ for the 2D simulation), $Fr = 0.512$ (runs 6 and 7), where $\mathrm{d}z = \pi/512k_z^{-1}$ is the grid cell of the two-dimensional simulation. Three-dimensional results are represented in gray and two-dimensional results in black. Theoretical prediction of the Hayes et al.'s model is represented with a dashed line (with $\mu = 37k_z = 0.2\lambda_0^{-1}$ in (A) and (B) and $\mu = 8.3kz = 0.2\lambda_0^{-1}$ in (C) and (D)).

gray patches in Fig. 5C). In fact, the existence of the sheet is correlated with breaking events: intense vortices associated with wavebreaking "dig" a local step in the density profile. For standing waves, these most intense events occur near the edge of the domain. However, similar association of a mixing vortex and a stratified sheet is also obtained for propagating waves in a periodic domain (e.g. Fig. 12 from Bouruet-Aubertot et al., 2004). The sheet has a wide extent in the y direction normal to the primary wave (Figs. 5B and D). We, therefore, conclude that the shape of these sheets is largely controlled by the dominant primary wave, even in the presence of significant 3D perturbations.

Strongly stratified sheets are still enhanced when the amplitude of the wave is increased (see Fig. 6), the extremum vertical density gradient is significantly stronger than in the previous case (−26 in Fig. 6D to be compared with −15 in Fig. 5D).

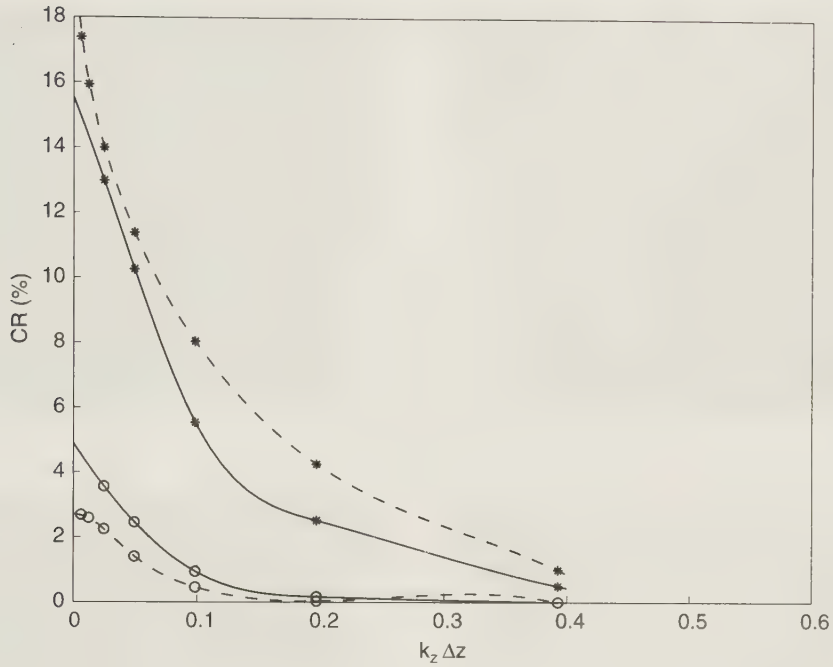


Fig. 4. Proportion of convective regions as a function of the vertical scale, for two values of the Froude number: $Fr = 0.512$ (filled circles, runs 6 and 7) and $Fr = 0.256$ (empty circles, runs 1 and 2). The 2D calculations are represented with a dashed line and the 3D calculations with a continuous line. Only four curves are represented for clarity.

Convective instabilities are then clearly seen in the transverse direction (Fig. 6B), therefore overturning events are more efficient in building up the step, as previously noticed in Figs. 3C and D.

The vertical extent of the sheet is small, about a few Ozmidov length scales: 7.26×10^{-2} for the two cases in Figs. 5 and 6, i.e. $5\lambda_0$ for $Fr = 0.512$ and $12\lambda_0$ for $Fr = 0.256$. In order to investigate the persistence of these structures we have introduced a forcing such that the energy of the primary wave is kept constant. Sheets result from independent wavebreaking events and their persistence is limited in time, lasting for a few buoyancy periods (see Bouruet-Aubertot et al., 2004).

In summary, the density gradients are strongly intermittent at small vertical scales, with a strong asymmetry between stable and convective regions. The intermittency results from breaking events that modify the vertical density structure into a succession of turbulent convective regions and strongly stratified sheets. The main question we address in the next section is whether any information on mixing can be inferred from these pdfs.

5. Link with vertical mixing

5.1. A model for the density gradient histograms

Two processes may be responsible for the increased skewness of the pdf of density gradients at small vertical scale, either the reversible straining produced by internal gravity waves or irreversible mixing events. These two alternatives were addressed in theoretical models. The effect of reversible straining has been considered by Desaubies and Gregg (1981). Assuming Gaussian statistics for the vertical wave displacement, they find a non Gaussian statistics for the difference of displacements. It is more and more skewed as the distance is reduced, as we expect for intermittency. The second model is the gamma pdf, proposed by Hayes et al. (1975), who assume a staircase density

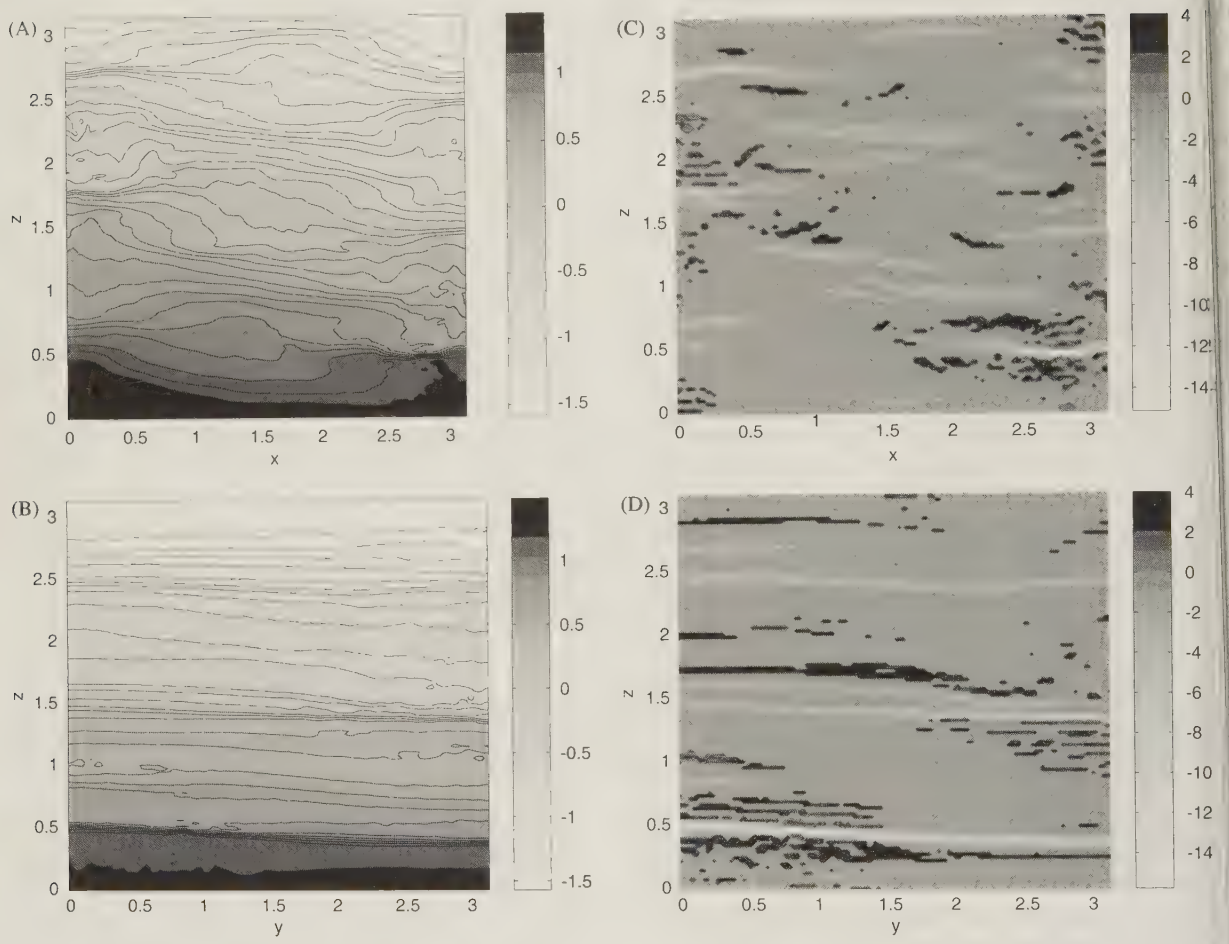


Fig. 5. Strongly stratified sheets with an initial dominant primary wave $\vec{k} = (1, 0, 1)$, $Fr = 0.256$ (run 1 in Table 1). The density fields (top) and vertical density gradients (bottom) are represented in two perpendicular vertical planes crossing at the location of extremal vertical density gradient. (A) Density field in the (x, z) plane at $y = 0.83$. (B) Density field in the (y, z) plane at $x = 2.79$. (C) Vertical density gradient in the (x, z) plane at $y = 0.83$. (D) Vertical density gradient in the (y, z) plane at $x = 2.79$. We notice that the strongly stratified sheets (white) are surrounded by convective zones (unstable density gradients represented in dark gray).

profile with randomly distributed steps separating mixed layers with uniform density. This model provides a good fit to oceanic data in the range of vertical scales (1–10 m), but it clearly implies excessive mixing. However the gamma pdf, obtained as a sum of statistical variables with Poisson statistics, can be justified in more general terms. It is infinitely divisible, meaning that a sum of independent random variables with this statistics has the same statistics (like for a Gaussian). Later Pinkel and Anderson (1992) considered a more realistic density profile since they allow partial mixing between sheets, assuming layers with random gradients. Interestingly, because of the above property of the gamma pdf, the pdf model for density differences has the same analytical expression (Eq. (16)) than that predicted by Hayes et al. (1975).

This model provides also a reasonable fit to our numerical data at scales smaller than $30\lambda_0$ (see Bouruet-Aubertot et al., 2004). It does not match with the weakly nonlinear regime at larger scales,

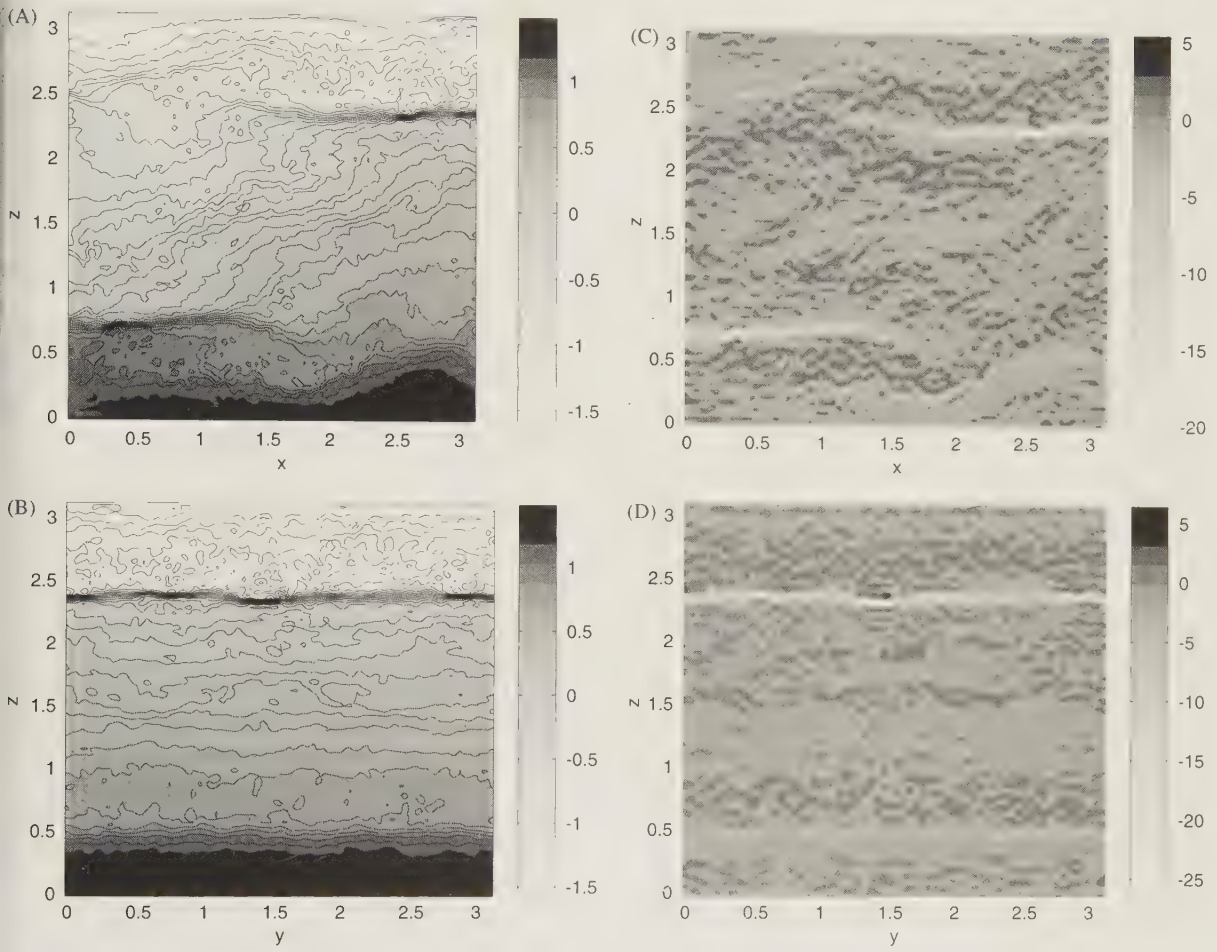


Fig. 6. Strongly stratified sheets in a three-dimensional simulation, with a primary wave of high amplitude $Fr = 0.512$ (run6 in Table 1). The density fields (top) and vertical density gradients (bottom) are represented in two perpendicular vertical planes crossing at the location of extremal vertical density gradient. (A) Density field in the (x, z) plane at $y = 1.32$. (B) Density field in the (y, z) plane at $x = 2.57$. (C) Vertical density gradient in the (x, z) plane at $y = 1.32$. (D) Vertical density gradient in the (y, z) plane at $x = 2.57$. We notice that the strongly stratified sheets (white) are surrounded by convective zones (unstable density gradients represented in dark gray).

for which the model of Desaubies and Gregg (1981) is found more appropriate. However mixing effects are controlled by the smaller scales, for which the gamma pdf model is appropriate. The probability density of the density difference $\Delta\rho$ at separation Δz is given by (see Eq. (9) of Hayes et al.)

$$P_g(\Delta\rho \mid \Delta z) = \frac{\mu}{\beta\Gamma(\mu\Delta z)}\left(\mu\frac{\Delta\rho}{\beta}\right)^{\mu\Delta z-1}\exp\left(-\mu\frac{\Delta\rho}{\beta}\right),$$

where Γ represents the Gamma function and β the mean density gradient (equal to -1 in our computations). The distribution at all scales Δz is determined by the single parameter μ. In the model of Hayes et al. (1975) μ denotes the mean number of steps per unit of height.

One easily checks that the integral of this probability density is unity and that the moment of degree n is $\langle(\Delta\rho)^n\rangle = \int_0^\infty P(\Delta\rho)(\Delta\rho)^n d(\Delta\rho) = (\beta/\mu)^n(\mu\Delta z)(\mu\Delta z+1)\cdots(\mu\Delta z+n-1)$. In particular the mean $n = 1$ is equal to the mean gradient β,

as it should. The variance of the fluctuations is $\langle(\Delta\rho - \langle\Delta\rho\rangle)^2\rangle = (\beta/\mu)^2\mu\Delta z$. The density fluctuations can be considered as resulting from a "random walk" with steps β/μ. The variance is proportional to the mean number of steps $\mu\Delta z$. Note that the third-order moment for density fluctuations varies as well linearly with the vertical separation: $\langle(\Delta\rho - \langle\Delta\rho\rangle)^3\rangle = 2(\beta/\mu)^3\mu\Delta z$. These relations provide a first test of the model, and a fit of its parameter μ. We have plotted these two moments as a function of the vertical distance Δz in Fig. 7. It can be reasonably fitted by a linear relationship in a limited range. The mean distance between sheets μ^{-1}, is inferred from the corresponding slopes ($\beta = -1$ in our computations).

Once the single parameter μ has been determined, the probability distribution at different scales is determined by the model, which provides a severe test (see Fig. 16 Bouruet-Aubertot et al., 2004 for comparison). The comparison shown in Fig. 3 is fairly good for vertical separations $\Delta z \approx [\lambda_0, 30\lambda_0]$. Nevertheless, the tails of the gamma pdf model tend to an exponential at small separations, while flatter tails are observed: intermittency is underestimated by the model (Figs. 3A and D). Furthermore, the occurrence of convective zones with positive $\Delta\rho$ is ignored.

5.2. An indirect estimate of energy dissipation and mixing

The formation of sheets and layers is clearly associated with energy dissipation and vertical mixing. We, therefore, expect a link between the Ozmidov scale, related to energy dissipation, and the mean vertical distance between sheets, μ^{-1}. Fig. 8 shows indeed that these two quantities are proportional $\mu^{-1} \approx 5.3\lambda_0$.

We next computed the vertical eddy diffusivity K_d using the method proposed by Winters and D'Asaro (1996) that leads to the following expression for the Cox number:

$$Cox = \frac{K_d}{\kappa} = \frac{\langle\|\nabla\rho\|^2\rangle_I}{(d\rho_b/dz)^2},$$

where $\langle\rangle_I$ denotes an average over an isopycnal surface and ρ_b refers to the sorted density profile. The key difference with the Osborn and Cox (1972) expression for the Cox number is that the average of the squared density gradient is computed over an isopycnal surface in place of a horizontal plane and the vertical gradient of the sorted density profile is taken into account in place of the horizontal mean of the vertical density

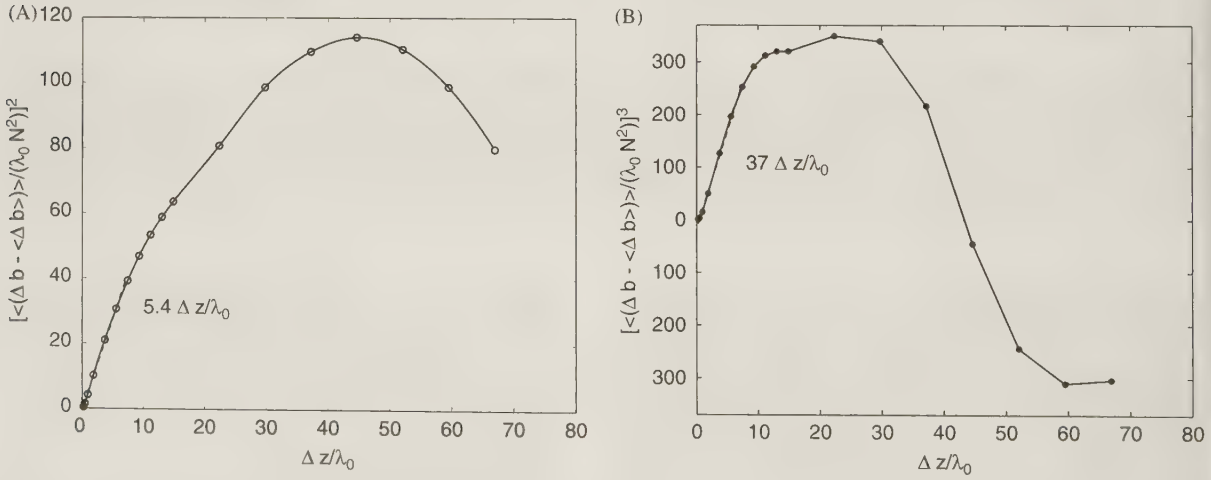


Fig. 7. (A) Variance of reduced density differences $\Delta b = g/\rho_0(\rho(x, z + \Delta z) - \rho(x, z))$ (in units of $(\lambda_0 N^2)$ as a function of vertical separation Δz (in units of λ_0) (run7), the slope at small scales is equal to β^2/μ (see text), i.e. $1/(\mu\lambda_0)$ in dimensionless form. (B) $\langle[\Delta b - \langle\Delta b\rangle]^3\rangle$ (in units of $(\lambda_0 N^2)^3$) as a function of Δz (in units of λ_0) (run7), the slope at small scale is equal to $2\beta^3/\mu^2$ (see text), i.e. $2/(\mu\lambda_0)^2$. The sheet spacing μ^{-1} is inferred from a linear fit of the two curves for vertical separation within $[0.2\lambda_0, 7.5\lambda_0]$, a value of $5.4\lambda_0$ is obtained.

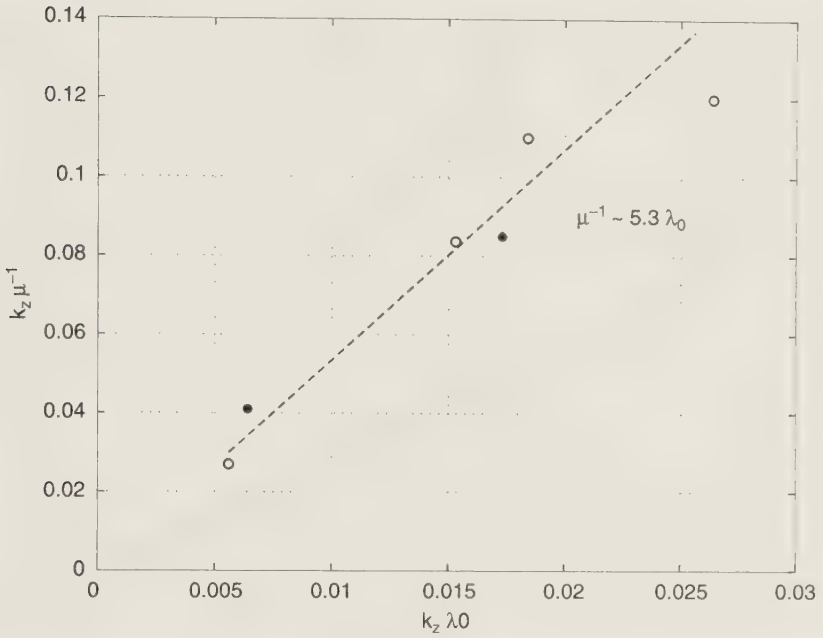


Fig. 8. Mean distance between steps (μ^{-1}) as a function of the Ozmidov length scale ($\lambda_0 = (\varepsilon/N^3)^{1/2}$). 2D simulations are represented with empty circles and 3D simulations with filled circles.

gradient. The drawback of the latter method is that it overestimates the Cox number when the mean vertical density gradient becomes small. We found indeed that while an excellent agreement between the two methods was obtained for the smallest wave amplitude, a discrepancy of a factor of 2 appears for the highest wave amplitude. Note also that in the method of Thorpe, a sorting of the fluid parcels is used along a vertical line of measurement, so it can be viewed as a simplified version of the method of Winters and D'Asaro suitable for field data.

The vertical eddy diffusivity K_d associated with wave breaking is classically related to energy dissipation, following Osborn (1980). Indeed the vertical mixing requires an increase of potential energy, to lift the heavier fluid. Part of the wave energy decrease corresponds to a conversion to this background potential energy. This results in a relationship

$$K_d = \frac{\gamma}{\gamma + 1}\frac{\varepsilon}{N^2},$$

where $\gamma/(\gamma+1) \leqslant 1$ is a mixing efficiency. We plot K_d/κ versus $\varepsilon/\upsilon N^2$ in Fig. 9. The two quantities are indeed proportional, with a similar result for two- and three-dimensional computations. The slope determines a high mixing efficiency $\gamma/(\gamma+1) = 0.7$ which deserves further comments. Indeed estimates in the ocean and in some laboratory experiments lead to a smaller mixing efficiency, about 0.2. Nevertheless, our result is more consistent with that inferred from laboratory experiments by Barry et al. (2001) who found that

$$\frac{K_d}{\kappa^{1/3}\upsilon^{2/3}} = 0.65\frac{\varepsilon_{ke}}{\upsilon N^2}$$

for turbulence intensity ($\varepsilon_{ke}/\upsilon N^2$), within the interval [10,300]. This agrees with our simulations at Prandtl number $Pr = (\upsilon/\kappa) = 1$, but yields lower values for smaller Pr. Indeed mixing efficiency in similar computations was found to decrease with increasing Prandtl number, reaching a value of 0.3 for $Pr = 10$ (Bouruet-Aubertot et al., 2001). In laboratory experiments, using salt concentration, $Pr = 1000$, so that a still smaller

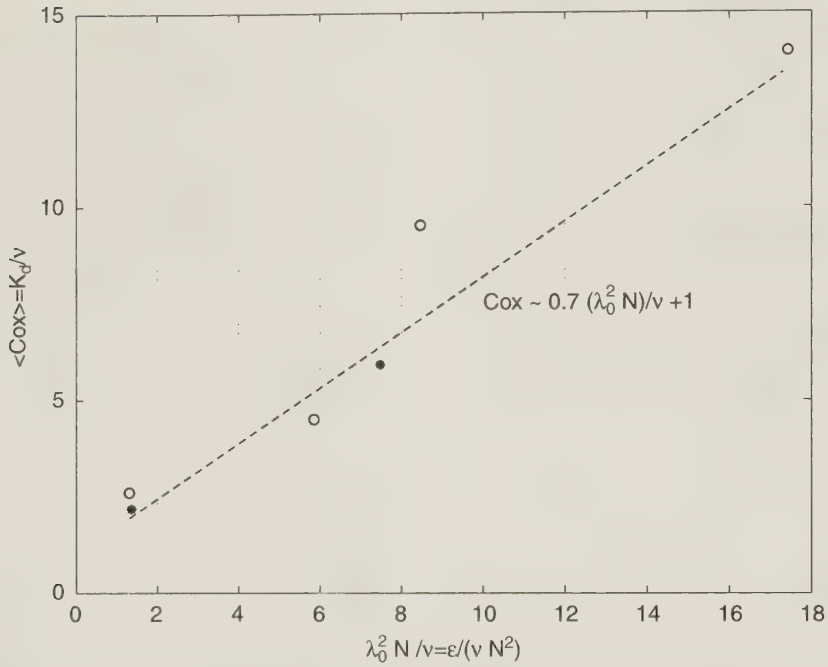


Fig. 9. Mean Cox number as a function of dimensionless total energy dissipation $\lambda_0^2 N/\nu$, 2D simulations are represented with empty circles and 3D simulations with filled circles.

mixing efficiency is expected. We expect that such Prandtl number dependency is most important at the lowest values of the turbulent intensity.

Combining this result with the link between λ_0 and μ^{-1}, we can relate the vertical eddy diffusivity with μ. In the next section we shall perform a similar analysis of the vertical density gradients to oceanic data and deduce estimates of energy dissipation ε and vertical eddy diffusivity K_d.

6. Application to in situ oceanic measurements in an area of energetic internal tides

We have analyzed a set of 24 temperature and salinity profiles collected during the ARCANE experiment in 1996 at a fixed point, (42° 06′N, 9° 33′W), located on the continental slope west of the Iberian peninsula (Le Cann et al., 1998). These profiles were acquired every hour during one day and the effective vertical resolution is equal to 1 m. A temperature salinity diagram of these data is shown in Fig. 10. Surface waters are characterized by temperatures greater than 13.5° and a mean salinity of 35.9 psu. Deeper, North-Atlantic central water (NACW) is represented by [AB] while point C corresponds to Mediterranean water (MW). The line BC results from mixing between MW and some water mass formed by a subduction process further North (point B). Point D represents Labrador sea deep water (LSW) and the region between C and D is characterized by salt-fingering processes.

The corresponding mean profiles of potential density σ, temperature T and salinity S, are represented in Fig. 11. Below the thermocline the density profile is almost linear down to about −1200 m following which the stratification is weaker. In the interior region, for depth within [−1000 m, −400 m], the temperature is almost constant (Fig. 10B) and the stratification is due mostly to salinity in this region (Fig. 11C). We restrict our analysis in this depth interval.

In this region the dynamics is characterized by the presence of a strong baroclinic M2 tide

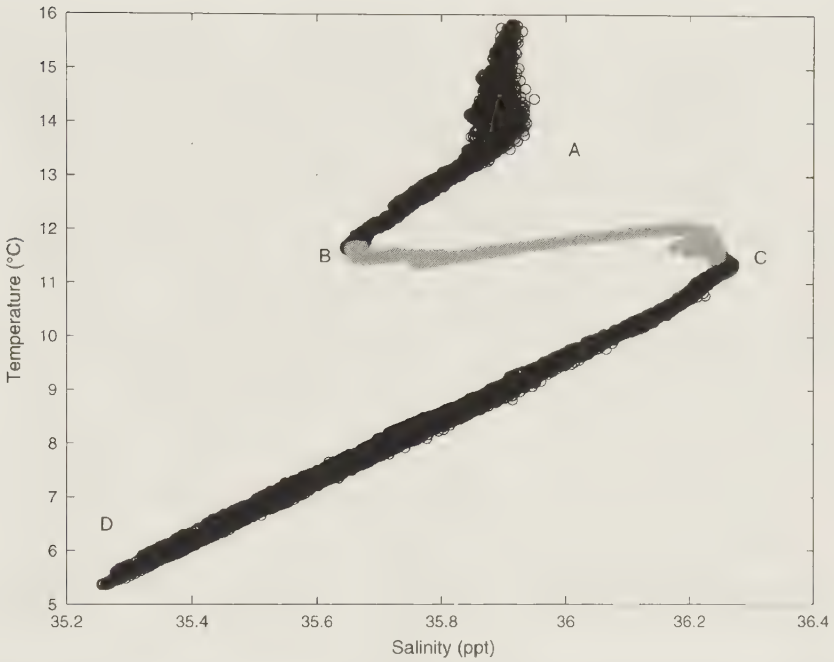


Fig. 10. Temperature salinity diagram taken from repeated CTD profiles at (42°06′N, 9 33′W) west of the Iberian peninsula during the ARCANE experiment. The analysis is focused on the data points represented in grey (red in online version) and corresponding to depth within [−1000 m, −400 m].

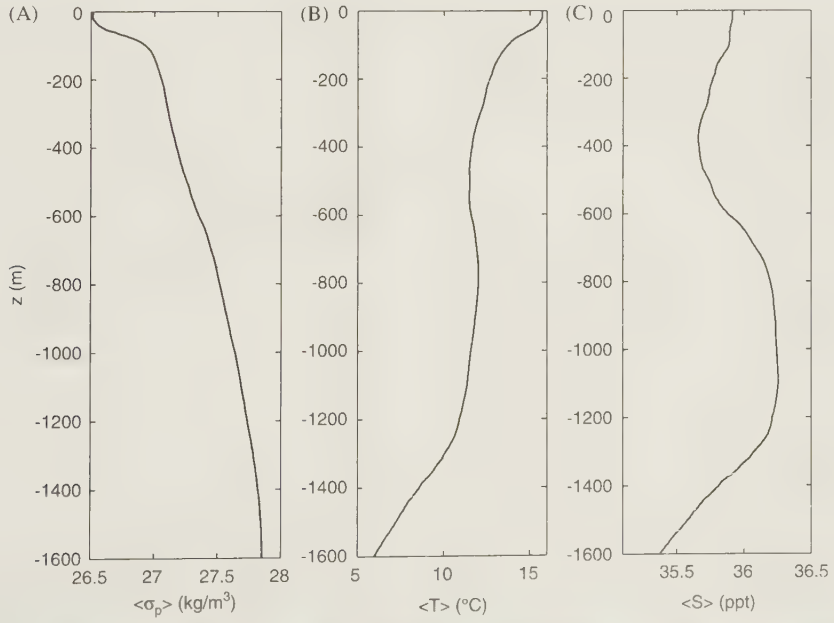


Fig. 11. Temporal mean of the vertical profiles: (A) potential density σ_p, (B) temperature T, (C) salinity S. Variances of σ_p, T and S are, respectively, equal to 2.3×10^{-4} (kg/m^3)^2, 3.0×10^{-3} (°C)^2 and 5.5×10^{-4} for depth within [−1000 m, −400 m].

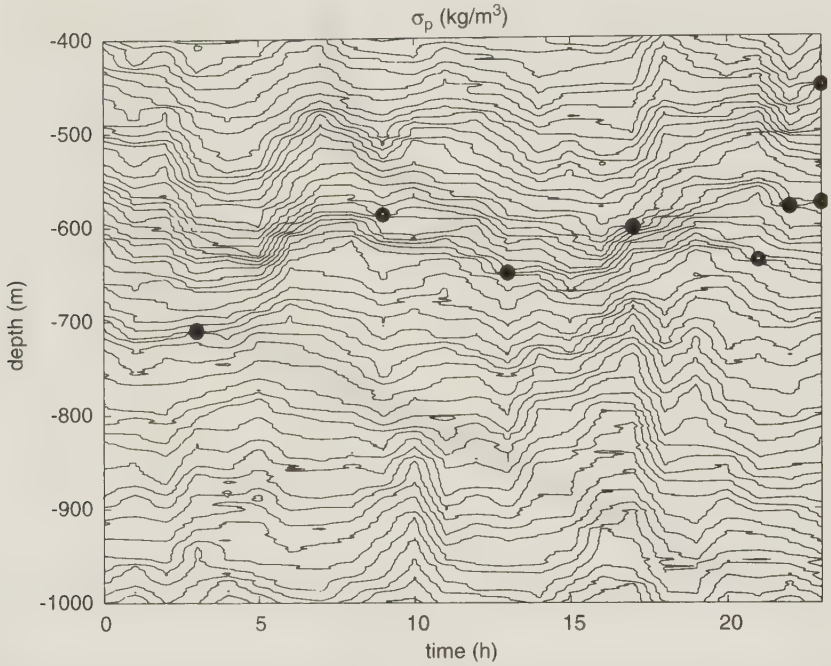


Fig. 12. Time depth map of potential density inferred from the 24 CTD profiles for depth within [−1000 m, −400 m], most stratified events are indicated with bold circles.

as shown on time depth section of the density (see Fig. 12, especially wave crests between [−1000 m, −400 m]). Furthermore, we checked that double diffusion can be dismissed since the stability condition is verified (Turner, 1973):

$$\frac{Ra}{Ra_s} \approx 0.17 < \frac{Pr+\tau}{1+Pr} = 0.87,$$

where Ra and Ra_s are the Rayleigh and saline Rayleigh number and τ the ratio between saline and thermal diffusivities. The mean buoyancy frequency is equal to 2.7×10^{-3} rad/s. Over this region the mean density variance is equal to 2.3×10^{-4} (kg/m^3)^2 which is much larger than the experimental error of about 5×10^{-6}.

We first investigated the vertical density structure to search for intermittency. Fig. 13 shows pdf of density differences computed for different vertical separations. Similarly to the numerical simulations, there is an increasing asymmetry as the vertical separation decreases while the width of the distribution increases. This asymmetry is correlated with the occurrence of overturning events ($\Delta\rho > 0$) and is observed for vertical separations smaller than 32 m. Comparison with the pdf of temperature and salinity differences confirms that this strong skewness is associated with the occurrence of static instabilities in a stratified medium. Indeed, the pdfs of temperature differences are symmetric even at small vertical scales (Figs. 14A, C and E), resulting from the fact that the mean vertical gradient is very small (1.14×10^{-4} °C/m). In contrast, the pdfs of salinity differences that are strongly skewed at small scales correspond to a stable mean salinity profile.

We next identified the strongly stratified regions seen on the pdf at small scales (asymmetric negative values in Figs. 13A–C). To this aim we selected pairs of locations for which the density difference for $\Delta z = 4$ m is larger (in absolute value) than 7 times the mean square root. Nine events were thus obtained that are represented with bold circles in Fig. 12 on the potential density

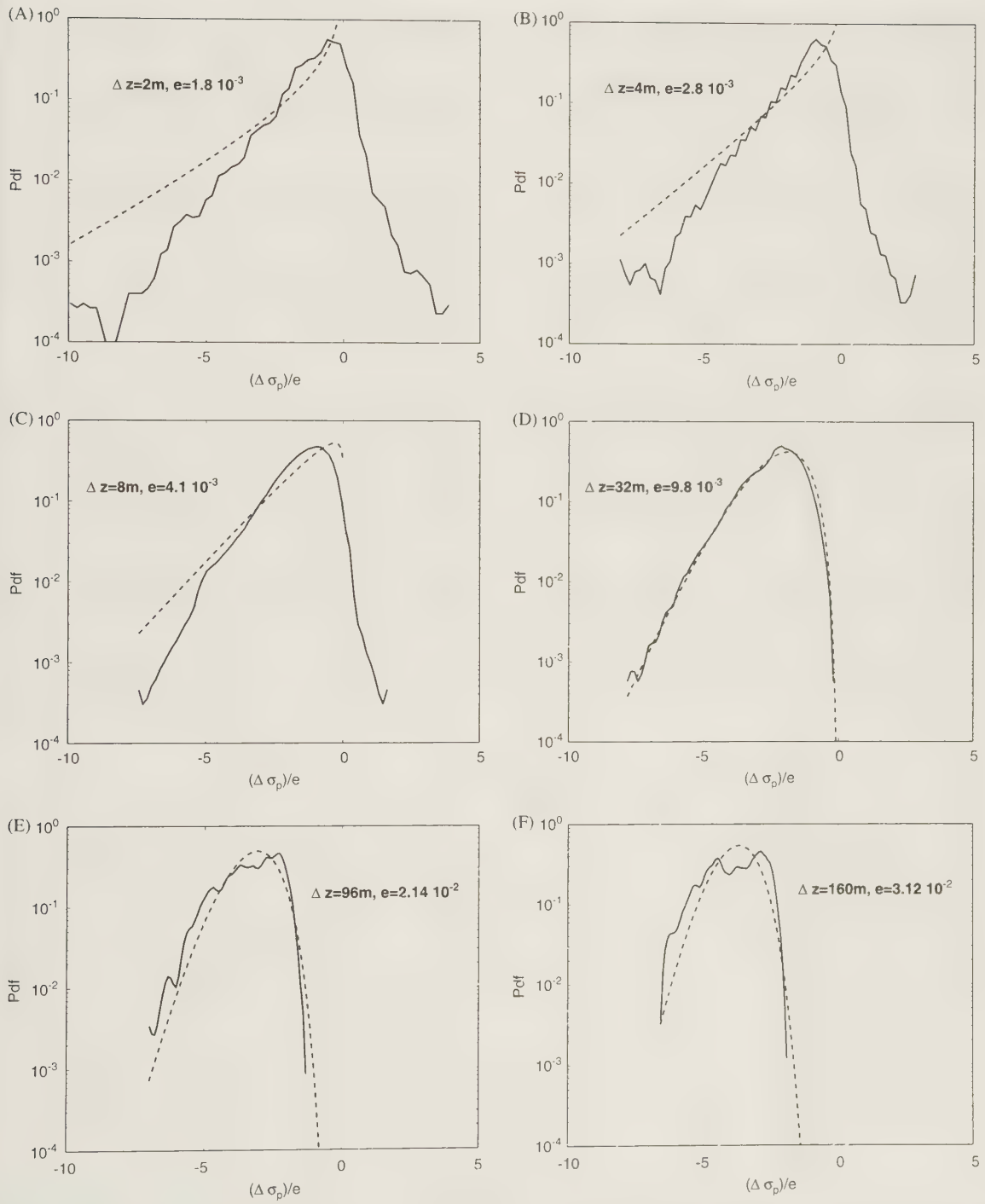


Fig. 13. Probability density functions of density differences for successive vertical separations Δz for depth within $[-1000\,\mathrm{m}, -400\,\mathrm{m}]$, normalized at each scale by the standard deviation e (given in $\mathrm{kg/m^3}$). Prediction of the Hayes et al. model with parameter $\mu = 0.16\,\mathrm{m^{-1}}$ (see text) is represented with a dashed line. (A) $\Delta z = 2\,\mathrm{m}$, (B) $\Delta z = 4\,\mathrm{m}$, (C) $\Delta z = 8\,\mathrm{m}$, (D) $\Delta z = 32\,\mathrm{m}$, (E) $\Delta z = 96\,\mathrm{m}$, (F) $\Delta z = 160\,\mathrm{m}$.

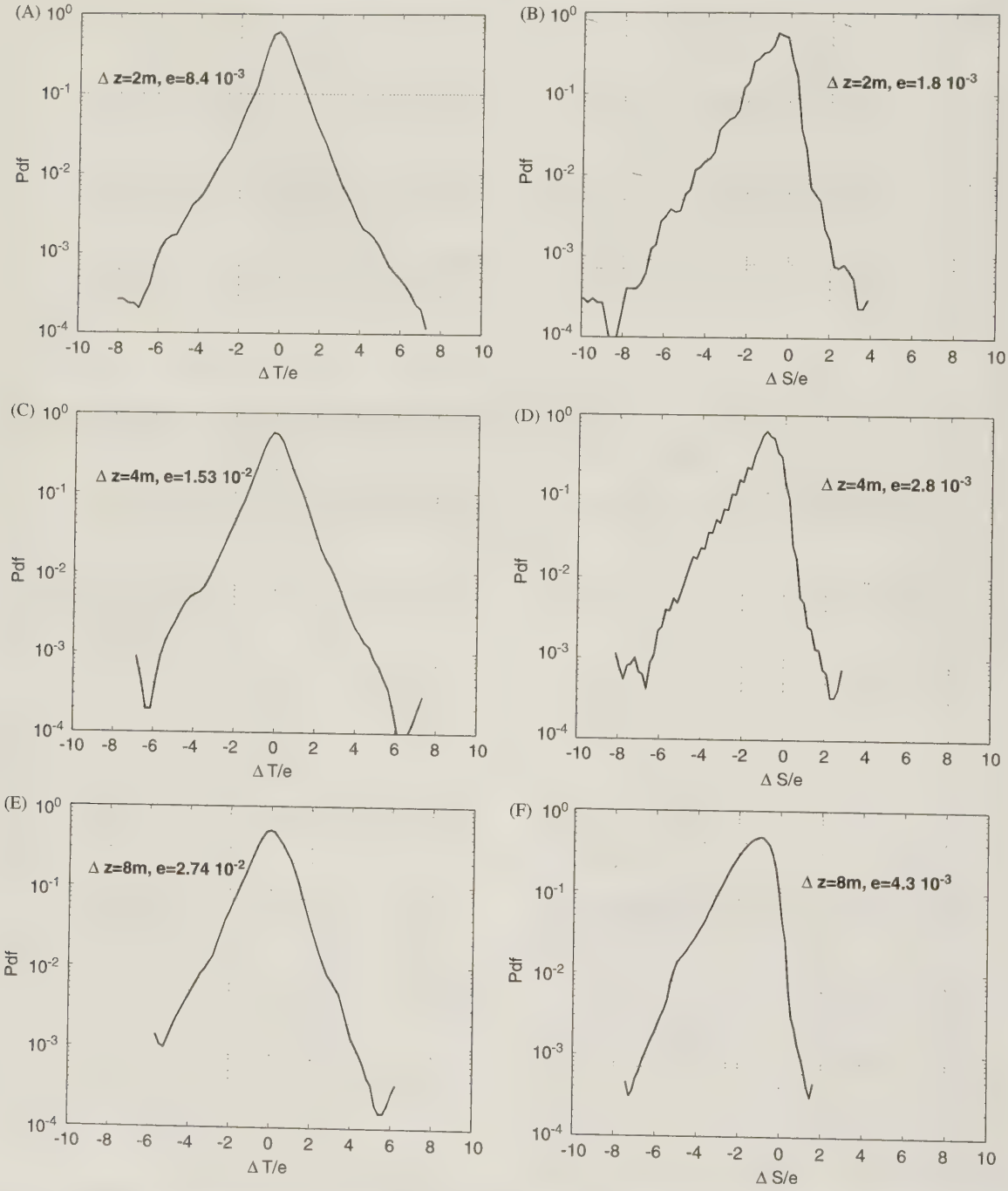


Fig. 14. Probability density functions of temperature and salinity differences for successive vertical separations Δz for depth within [−1000 m, −400 m], normalized at each scale by the standard deviation e (given in °C and in ppt, respectively, for T and S). (A) pdf(ΔT) at $\Delta z = 2$ m, (B) pdf(ΔS) at $\Delta z = 2$ m, (C) pdf(ΔT) at $\Delta z = 4$ m, (D) pdf(ΔS) at $\Delta z = 4$ m, (E) pdf(ΔT) at $\Delta z = 8$ m, (F) pdf(ΔS) at $\Delta z = 8$ m.

map. Profiles of the vertical density gradient plotted at the time of occurrence of these events give better evidence of the density steps, three of them are represented in Figs. 15A–C. The four main density steps are indicated by arrows in these figures.

A zoom on the event occurring at $t = 13$ h is shown in Fig. 16. A strong density step is observed while overturning events occur on both sides, in a similar way as in the numerical simulations, thus building up the step. Note that lateral thermohaline intrusions do not contribute significantly to this process since a linear relationship between T and S is obtained (Fig. 16B). Zooms on salinity and temperature profiles give evidence of their relative contribution to density variations. The variations of salinity leading to a density overturning play here a major role in building up the density sheet. Indeed, a strong salinity step is observed, that is able to over-compensate the inverse temperature step producing a strong density sheet. Moreover, an inverse temperature sheet is generated by convergent motions produced by breaking events in an unstable background temperature profile.

We have tested the Hayes et al. model against our oceanic data. We proceeded in a similar way as in the numerical simulations. The mean density gradient is inferred from the evolution of mean density differences as a function of vertical separation (Fig. 17A) and the parameter μ was inferred from the variance of density differences (Fig. 17B). Note that the linear fit was performed in an interval that is consistent with that of the observed skewness of the pdf: $\Delta z \approx [2\,\mathrm{m}, 32\,\mathrm{m}]$. A value of 5.9 m is obtained for μ^{-1}, which using $\mu^{-1} \approx 5.3\lambda_0$, corresponds to an Ozmidov scale, as previously defined, of 1.1 m, and a total energy dissipation of $\varepsilon = 2.4 \times 10^{-8}\,\mathrm{W/kg}$. If we further assume a mixing efficiency of 0.2 (Oakey, 1982) this leads to an estimate of the mean eddy diffusivity $K_\mathrm{d} = (\gamma/(\gamma+1))\lambda_0^2 N = 5 \times 10^{-4}\,\mathrm{m^2\,s^{-1}}$. This fairly high value, about 50 times larger than the estimated background value of 10^{-5}, agrees within a factor of 2 with the eddy diffusivity inferred from the Thorpe scale following Dillon (1982) and is consistent with the presence of strong internal tides (e.g. Lien and Gregg, 2001).

The shape of the pdfs agrees well with the model within the range $\Delta z \approx [4\,\mathrm{m}, 32\,\mathrm{m}]$, as shown in

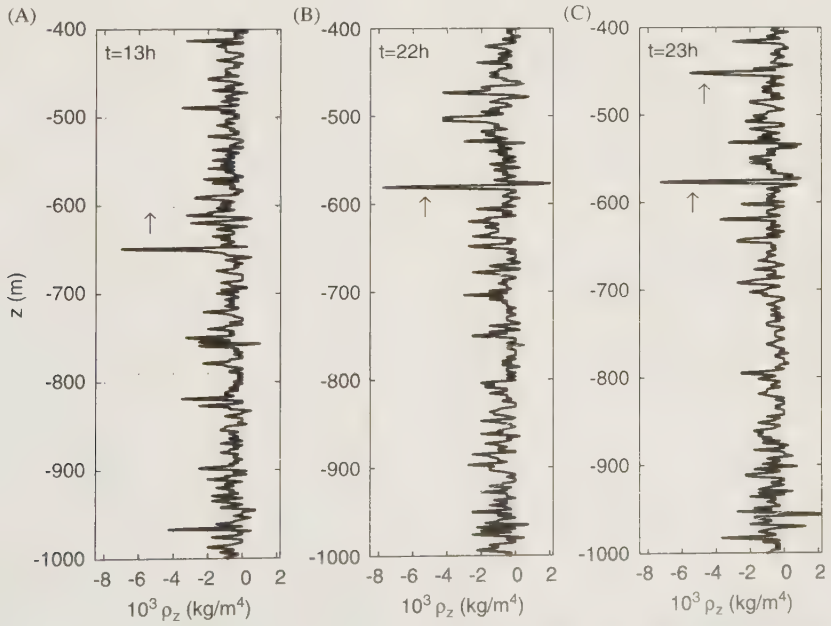


Fig. 15. (A), (B) and (C): Profiles of the vertical density gradient (at 2 m) at a given time showing some of the most intense events (indicated by arrows): (A) at time $t = 13$ h, (B) at time $t = 22$ h, (C) at time $t = 23$ h.

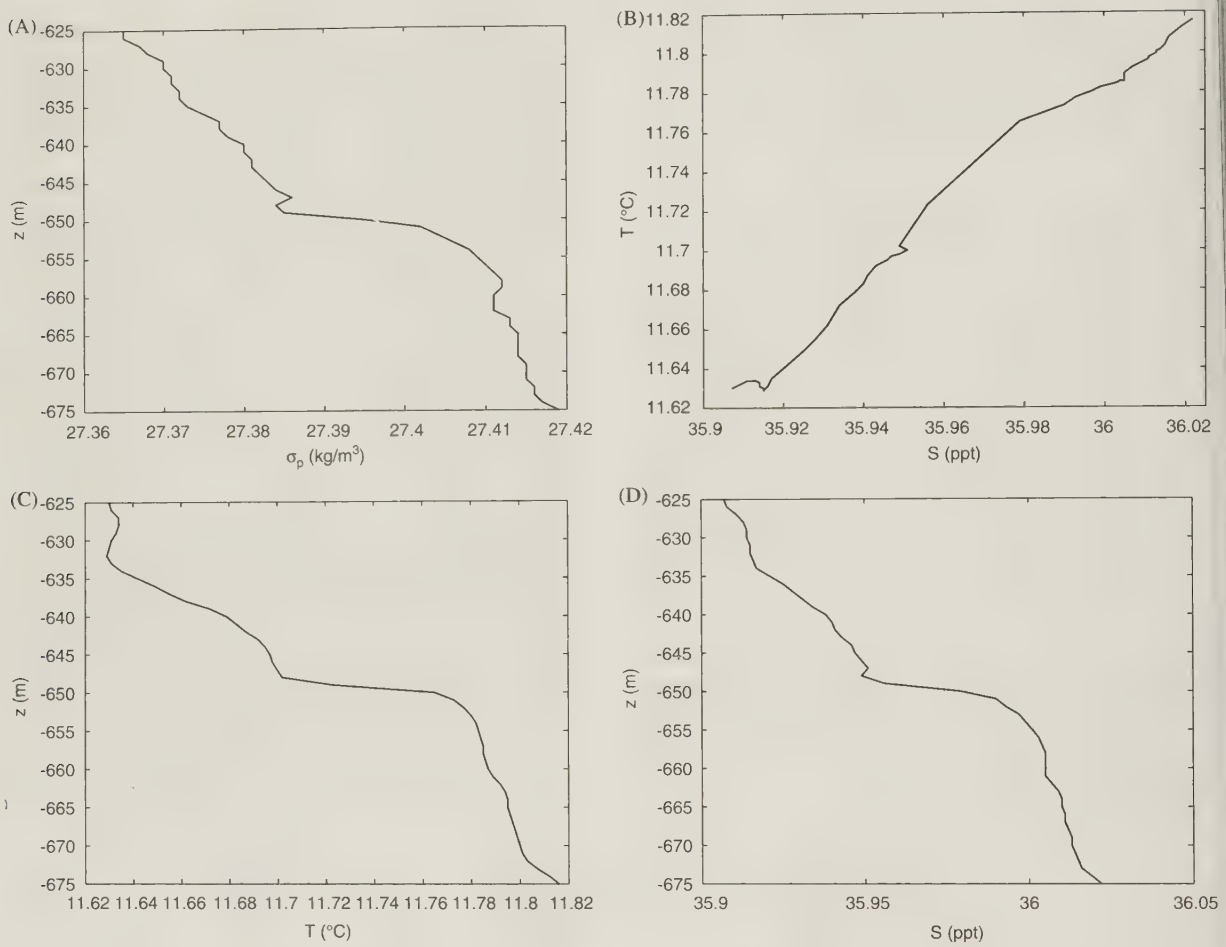


Fig. 16. Zoom over one intense event occurring at $t = 13$ h: (A) potential density profile σ_p, (B) vertical potential density gradient profile $\partial_z \sigma_p$, (C) temperature profile T, (D) salinity profile, S.

Fig. 13. Also, the percentage of convective regions decreases rapidly as the vertical separation increases. A maximum of 7% is obtained, that is within the range of our numerical results (see Fig. 4).

7. Conclusion

We analyzed some statistical properties of the turbulence induced by the breaking of internal gravity waves. This dynamics corresponds to the so-called buoyancy subrange and it is characterized by kinetic and available potential energy spectra that follow a k_z^{-3} power law. In the ocean such behaviour is obtained in the finescale domain, in between weakly nonlinear waves and three-dimensional turbulence, corresponding to vertical scales typically within the range [1 m, 10 m]. We focused on the vertical structure of the density field with the main purpose of providing a link between density intermittency, energy dissipation and mixing processes.

Density gradients are strongly intermittent at small scales. The main feature is the strong skewness of the histograms toward stable stratification that becomes evident as soon as overturning events occur. We have characterized these rare and intense events. These take the form of strongly stratified

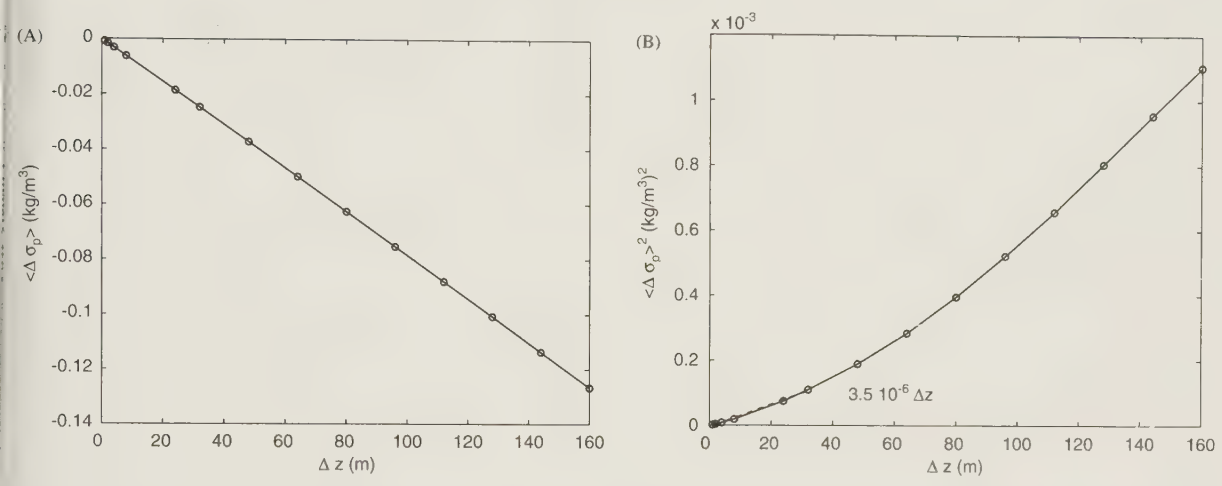


Fig. 17. (A) Mean of $\Delta\ \sigma_p$ as a function of the vertical separation, Δz; the slope β is equal to 7.84×10^{-4}, (B) Variance of $\Delta\ \sigma_p$ as a function of vertical separation Δz, the slope p within [2 m, 32 m] is equal to 3.5×10^{-6} and μ^{-1} is inferred from this linear fit $\mu^{-1} = p/\beta^2 = 5.9$ m.

sheets of thin vertical thickness close to the Ozmidov scale and horizontal scale related to that of the large scale breaking wave. These structures are generated through strong convergent motions induced by surrounding overturning events. Indications of these structures were obtained in the atmosphere (e.g. Dalaudier et al., 1994; Alisse and Sidi, 2000) from vertical profiles of temperature, where large stable temperature steps are observed (see Fig. 2 from Alisse and Sidi, 2000). However, because of advection by the mean wind, balloons could not track these sheets. In the ocean the same events seem to be observed on temperature cross-sections obtained by Shen and Mied (1986) in the seasonal thermocline of the Sargasso Sea.

We then checked the ability of the sheet and layer model or gamma pdf model (e.g. Hayes et al., 1975; Pinkel and Anderson, 1992) to reproduce the shape of the histograms of vertical density gradient at small vertical scales (typically within $\lceil \lambda_0, 30\lambda_0 \rceil$). This model, which assumes that the step-like structure is the result of mixing processes, depends on one main parameter, the mean number of sheets (or steps) per unit length, μ. This parameter μ or alternatively the sheet spacing μ^{-1} is inferred from the variance and third order moment of density differences over vertical scales within the domain of strong skewness of the pdfs.

The main result of this paper is that the sheet spacing μ^{-1}, varies linearly with the Ozmidov length scale defined with the total energy dissipation rate. Thus this provides a link between density statistics at finescale and energy dissipation. As a consequence, the mean eddy diffusivity that characterizes breaking events varies linearly with a dimensionless parameter defined with μ^{-1}: $\mu^{-2}N/v$. These findings motivated a preliminary comparison with oceanic data.

The analysis was next being applied to oceanic data from a region of strong internal tides west of the Iberian peninsula. We focused on a dynamically consistent region namely within the interior for depth between -1000 and -400 m. There, the mean vertical salinity and potential density profiles are stable and linear while the mean vertical temperature gradient is small and unstable in some regions. Thus, density differences are mainly due to salinity variations. Unsurprisingly histograms of density and salinity differences present an increasing skewness at small vertical scales while pdfs of temperature differences are symmetric even at small vertical scales. This underlines that a crucial ingredient to this skewness is the presence of a stable mean vertical gradient. Application of the sheet and layer model to these data gave fairly good predictions for vertical scales within

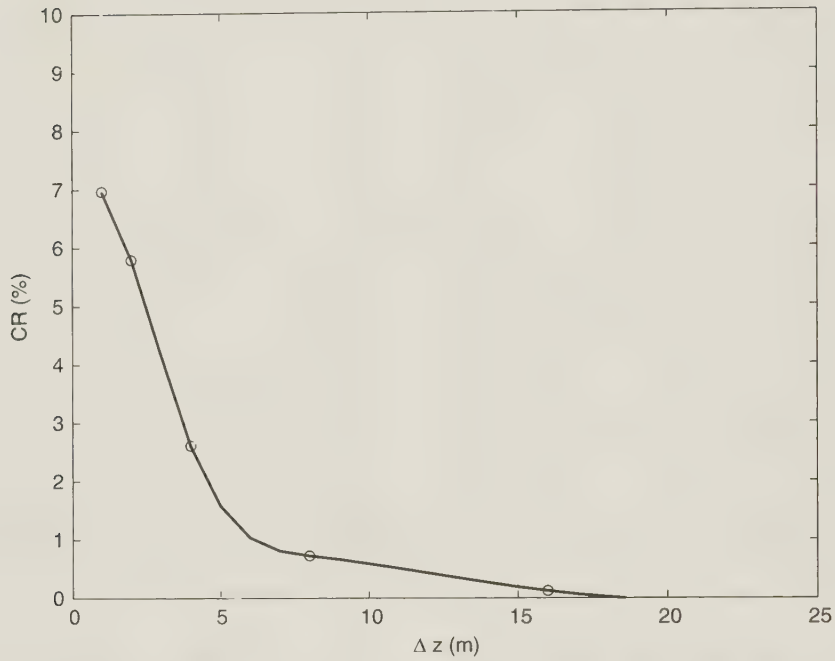


Fig. 18. Percentage of convective regions as a function of the vertical separation Δz.

[4 m, 32 m]. Finally we applied the relationship between μ^{-1} and the Ozmidov scale λ_0 obtained from the numerical simulations. This gave estimates of energy dissipation and eddy diffusivity consistent with common values obtained in regions of strong internal tides. In contrast if we apply our numerical results to the Hayes et al. results ($\mu^{-1} = 0.62$ m), an Ozmidov length scale of 12 cm and an eddy diffusivity of $1.5 \times 10^{-5}\,\mathrm{m^2/s^{-1}}$ are obtained. These values are relevant for an open ocean site remote from topographic features.

In future work we plan to further check this indirect estimate of mixing based on density statistics at finescale on a large set of in situ oceanic data and compare this estimate with direct measurements of mixing as inferred from microstructure measurements (Fig. 18).

Acknowledgements

The authors wish to thank F. Dalaudier for helpful discussions and N. Daniault for her comments regarding the analysis of the oceanic data. Numerical simulations were performed on the NEC SX-5 and the Cray T3E of the Institut du développement et des Ressources en Informatique Scientifique (IDRIS) under contracts 020580 and 980580. PBA acknowledges financial support from ACI CNRS "Jeunes chercheurs" No. 0693 and PATOM/CNRS.

References

Alisse, J.R., Sidi, C., 2000. Experimental probability density functions of small-scale fluctuations in the stably stratified atmosphere. Journal of Fluid Mechanics 402, 137–162.

Baker, M.A., Gibson, C.H., 1987. Sampling turbulence in the stratified ocean: statistical consequences of strong intermittency. Journal of Physical Oceanography 17, 1817–1836.

Barry, M.E., Ivey, G.N., Winters, K.B., Imberger, J., 2001. Measurements of diapycnal diffusivities in stratified fluids. Journal of Fluid Mechanics 442, 267–291.

Bénielli, D., Sommeria, J., 1998. Excitation and breaking of internal gravity waves by parametric instability. Journal of Fluid Mechanics 374, 117–144.

Bouruet-Aubertot, P., Sommeria, J., Staquet, C., 1995. Breaking of standing internal gravity waves through two-dimensional instabilities. Journal of Fluid Mechanics 285, 265–301.

Bouruet-Aubertot, P., Sommeria, J., Staquet, C., 1996. Stratified turbulence produced by internal wave breaking. Dynamics of Atmospheres and Oceans 23, 371–378.

Bouruet-Aubertot, P., Koudella, C.R., Staquet, C., Winters, K.B., 2001. Particle dispersion and mixing by breaking internal gravity waves. Dynamics of Atmospheres and Oceans 33, 95–134.

Bouruet-Aubertot, P., Sommeria, J., Koudella, C.R., 2004. Intermittency in stratified turbulence produced by breaking internal gravity waves. Journal of Fluid Mechanics, submitted for publication.

Canuto, C., Hussaini, M.Y., Quateroni, A., Zang, T.A., 1988. Spectral Methods in Fluid Dynamics, Springer Series in Computational Physics. Springer, Berlin.

Carnevale, G.F., Briscolini, M., Orlandi, P., 2001. Buoyancy to inertial range transition in forced stratified turbulence. Journal of Fluid Mechanics 427, 205–239.

Dalaudier, F., Sidi, C., Crochet, M., Vernin, J., 1994. Direct evidence of "sheets" in the atmospheric temperature field. Journal of Atmospheric Sciences 51, 237–248.

Desaubies, Y., Gregg, M.C., 1981. Reversible and irrreversible finestructure. Journal of Physical Oceanography 11, 541–556.

Dillon, T.M., 1982. Vertical overturns: a comparison of Thorpe and Ozmidov length scales. Journal of Geophysical Research 87, 9601–9613.

Gargett, A.E., Hendricks, P.J., Stanford, T.B., Osborn, T.R., Williams III, A.J., 1981. A composite spectrum of vertical shear in the ocean. Journal of Physical Oceanography 11, 1258–1271.

Gregg, M.C., D'Asaro, E.A., Shay, T.J., Larson, N., 1986. Observations of persistent mixing and near-inertial internal waves. Journal of Physical Oceanography 16, 856–885.

Gregg, M.C., Seim, H.E., Percival, D.B., 1993. Statistics of shear and turbulent dissipation profiles in random internal wave fields. Journal of Physical Oceanography 23, 1777–1799.

Hayes, S.P., Joyce, T.M., Millard, J.R., 1975. Measurements of vertical fine structure in the Sargasso Sea. Journal of Geophysical Research 80 (3), 314–319.

Klostermeyer, J., 1991. Two and three-dimensional parametric instabilities in finite amplitude internal gravity waves. Geophysical Astrophysical Fluid Dynamics 61 (1–4), 1–25.

Koudella, C.R., 1999. Ondes internes de gravité: instabilités, déferlement et vorticité potentielle. Thèse de doctorat, Ecole Normale Supérieure de Lyon.

Le Cann, B., Billant, A., Branellec, P., 1998. Campagne ARCANE 96, recueil de données: vol. 1: CTD-O2, Rapport interne LPO 98-01.

Lien, R.C., Gregg, M.C., 2001. Observations of turbulence in a tidal beam and across a coastal ridge. Journal of Geophysical Research 106 (C3), 4575–4591.

Lombard, P.N., Riley, J.J., 1996. Instability and breakdown of internal gravity waves: I. linear stability analysis. Physics of Fluids 8 (12), 3271–3287.

Oakey, N.S., 1982. Determination of the rate of dissipation of turbulent energy from simultaneous temperature and velocity shear microstructure measurements. Journal of Physical Oceanography 12, 256–271.

Olbers, D.J., 1976. Non-linear energy transfers and the energy balance of the internal wave field in the deep ocean. Journal of Fluid Mechanics 74, 375–399.

Osborn, T.R., 1980. Estimates of the local rate of vertical diffusion from dissipation measurements. Journal of Physical Oceanography 10, 83–89.

Osborn, T.R., Cox, C.S., 1972. Oceanic finestructure. Geophysical Fluid Dynamics 3, 321–345.

Pinkel, R., Anderson, S., 1992. Toward a statistical description of finescale strain in the thermocline. Journal of Physical Oceanography 22 (7), 773–795.

Pinkel, R., Anderson, S., 1997a. Shear, strain and Richardson number variations in the thermocline. Part I: statistical description. Journal of Physical Oceanography 27 (2), 264–281.

Pinkel, R., Anderson, S., 1997b. Shear, strain and Richardson number variations in the thermocline. Part II: modelling mixing. Journal of Physical Oceanography 27 (2), 282–290.

Turner, J.S., 1973. Buoyancy Effects in Fluids. Cambridge University Press, Cambridge.

Winters, K.B., D'Asaro, E.A., 1996. Diascalar flux and the rate of fluid mixing. Journal of Fluid Mechanics 317, 179–193.

Yamazaki, H., Lueck, R., 1990. Why oceanic dissipation rates are not lognormal. Journal of Physical Oceanography 20, 1907–1918.

ELSEVIER

Deep-Sea Research II 51 (2004) 2943–2971

DEEP-SEA RESEARCH
PART II

www.elsevier.com/locate/dsr2

Near-bed solibores over the continental slope in the Faeroe-Shetland Channel

Phil Hosegood*, Hans van Haren

Royal Netherlands Institute of Sea Research (NIOZ), P.O. Box 59, 1790 AB Den Burg, The Netherlands

Abstract

On two occasions within a 12-day measurement period in the Faeroe-Shetland Channel, strongly nonlinear wave trains were observed at the sea-bed, propagating *up* the continental slope in water depths >450 m. The events were separated by a period of 4 days and, whilst resembling in appearance a density current running up a slope, are termed 'solibores', displaying the properties of both turbulent internal bores and nonlinear internal solitary waves (ISW). Each solibore displays a steep leading edge followed by a train of nonlinear waves with amplitudes of $O(10\,\mathrm{m})$ and periods of $\sim$5–20 min. Wave-induced particle velocities are consistent with ISW, whilst a zone of strong horizontal convergence at the leading edge of the solibores causes the formation of a rotor with flow in the opposite sense to a 'forward overturning' surface wave. In both cases upward vertical velocities $>10\,\mathrm{cm\,s^{-1}}$ are immediately followed by a return downward flow of equal magnitude. A CTD and microstructure transect conducted during the passage of the first solibore illustrates its behaviour as an up-slope intrusion of cold, dense water with concurrent rates of turbulent dissipation, ε, of $O(10^{-7}\,\mathrm{W\,kg^{-1}})$ and resulting in short-term maximum vertical diffusivities, K_z, of $O(10^{-1}\,\mathrm{m^2\,s^{-1}})$. In the long-term however, $K_z<10^{-4}\,\mathrm{m^2\,s^{-1}}$, implying that the solibores are not important for sustaining deep-sea mixing. In contrast, sediment fluxes at 2 and 30 m above the bed during the first solibore are $O(10^2)$ larger than the background value, implying solibores are the dominant sediment transport mechanism despite their intermittent occurrence. The transport of sediment, up the slope in the direction of propagation of the solibore, is contrary to the usually considered pathway for sediment from coastal seas to the abyss via the continental slope. That sediment transport is dominated by short-term events appears to be corroborated by long-term data (>140 days) which again indicate intermittency in periods of enhanced sediment flux. The solibores are consistent with the results of numerical and laboratory experiments on hydraulic jumps resulting from shoaling ISW. The influence of the slope is proposed to be the reason for their slightly different forms and effects on sediment transport; thus the first solibore, propagating directly up the slope, is proposed to result from an overturning hydraulic jump caused by kinematic instabilities and which subsequently forms a horizontal density intrusion. The second solibore represents a dispersive wave train due its oblique direction of propagation which reduces the effective bottom-slope, allowing dispersion to balance nonlinear effects and prohibiting overturning.
© 2004 Elsevier Ltd. All rights reserved.

*Corresponding author. Tel.: +31 222 369406; fax: +31 222 319674.
E-mail address: hosegood@nioz.nl (P. Hosegood).

0967-0645/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dsr2.2004.09.016

1. Introduction

The continental slope strongly influences the hydrodynamics in its vicinity and has consequences for mixing processes and the transport of sediments and nutrients. The interaction of internal waves, especially their critical reflection (Eriksen, 1982), with the slope has traditionally been proposed to account for the enhanced mixing required to maintain the meridional overturning circulation (Munk, 1966; Munk and Wunsch, 1998), whilst internal tides promote mixing both at the slope where they are generated and reflected (New and Pingree, 1990), and in the ocean interior where they may propagate as distinct beams of energy (for example, Pingree et al., 1986) or interfacial waves propagating along the pycnocline (Holloway, 1987; Small et al., 1999). The latter case in particular is often observed in shelf seas as the degeneration of an internal tidal wave into an internal bore or a packet of solitons. In response to observations of such internal waves, Henyey and Hoering (1997) coined the term 'solibore' to define the high frequency internal waves which display the properties of turbulent internal bores whilst being accompanied by short-period intense pulse-like internal waves which display the characteristics of solitons or internal solitary waves (ISWs). The ISWs are of permanent form with a 'sech^2' profile, extending into the thicker of the two layers between which they propagate, with the consequence that they are usually observed as waves of depression given the location of the near surface thermocline. Internal bores and hydraulic jumps typically result from the nonlinear steepening of, in the ocean, an internal tidal wave or, in lakes, of a wind-forced basin-scale wave. They exhibit strong dissipation and shear, with the linear long wave phase speed behind the bore exceeding the speed of the bore itself (Henyey and Hoering, 1997).

The ambiguity in observations that prompted the definition of a solibore can arise due to the time at which observations are made. A solibore begins like a bore in which nonlinear steepening dominates, before becoming more wavelike as the energy supply is reduced and with it the amplitude and dissipation of the bore, with dispersion subsequently balancing the nonlinearity as in a train of solitons. Thus whether a solibore appears in observations more as a bore or a nonlinear wave train may be simply a consequence of at what stage of its evolution the measurements are made. Observations from shelf breaks (Holloway, 1987; Small et al., 1999; Colosi et al., 2001; Johnson et al., 2001) indicate the manifestation of the shoreward propagating internal tide as an internal bore. In each case, however, the bore is accompanied by, or evolves into, packets of solitary waves. Observations on the continental shelf proper suggest the form of the internal tidal wave to be a packet of nonlinear waves, or more specifically solitons (Sandstrom and Elliot, 1984; Howell and Brown, 1985; Jeans and Sherwin, 2001a, b) consistent with the concept of the solibore beginning more as a bore before evolving into a wave train. Solibores with vertical displacements in excess of 100 m are also common due to tidal forcing in the Straits of Gibraltar (Wesson and Gregg, 1988; Morozov et al., 2002) and are typically accompanied by short-period internal waves which have been described as solitons (Bryden et al., 1994).

Whilst a solibore was originally defined to describe the evolution of internal tides, the term may be extended to include features that demonstrate the same bore-like properties accompanied by high-frequency solitary waves but which are forced by motions other than the tide. Solibores may be forced by the wind in lakes where internal seiches and standing waves steepen during their propagation along the axis of the lake (Hunkins and Fliegel, 1973; Thorpe, 1977; Horn et al., 2001). Horn et al. (2001) found the evolution of a wind induced, interfacial gravity wave into either a train of solitons or a turbulent internal bore, accompanied by strong dissipation and mixing due to localised shear-induced turbulence, to be dependent on the initial amplitude of the wave. A larger initial wave tended towards the formation of a bore. ISWs observed near a steepened Kelvin wave front and at the head of an intrusive thermocline jet in Lake Biwa and Lake Kinneret potentially contain as much as 1% of the basin-scale Kelvin wave energy and represent an important pathway for the transferral of large-scale energy to small scales which are available for mixing as the ISWs shoal at the lake perimeter (Boegman et al., 2003).

Shoaling ISWs have been shown recently in numerical and laboratory experiments to result in solibore-like features. Vlasenko and Hutter (2002) study numerically the shoaling of intense ISWs of depression propagating from deep water onto the shelf. When conditions are conducive for breaking, nonlinearity causes the rear face of the ISW to steepen, forming a baroclinic bore, which eventually overturns due to kinematic instabilities to produce 'soliton-like waves of elevation'. Such waves were also observed by Wallace and Wilkinson (1988), Helfrich (1992) and Lamb (2002) and are also referred to as 'turbulent boluses' and 'solitary-like waves with trapped cores'. They are however not positive solitary waves but rather exhibit similar behaviour; they correspond qualitatively to the crests on a wave packet that would be caused by the topographic scattering of a solitary wave of depression incident on a slope-shelf topography into a dispersive wave packet with waves of elevation (Djordjevic and Redekopp, 1978; Helfrich et al., 1984). The scattering process is interrupted when the interface intersects the slope with the boluses forming rather than a wave packet. The overturned wave then behaves as an upstream propagating pulsating jet, which exhibits wave-like undulations on its upper edge and whose crests extend into the thicker overlying layer. In contrast, Lamb (2002) proposes that the waves which form the wave packet resulting from a breaking ISW are indeed consistent with genuine ISWs rather than turbulent boluses, with the eventual wave field composed of a large leading solitary wave followed by one or two smaller solitary waves and then a mode-one wave train, a mode-two wave and other higher-mode waves. In some respects the pulsating jet may resemble a density current propagating up a slope, with the dense fluid penetrating beneath the lighter fluid higher on the slope. Such currents have been observed in the laboratory, where the initial head is cut off from the dense flow resulting in the formation of a new head behind it due to the interaction between the current and waves set up by its advancement (Simpson, 1997). Under no circumstances however has such a current been observed to generate a nonlinear dispersive train, whereas for less steep slopes than that required for the overturning of an ISW, dispersion has time to balance nonlinearity and permit the development of a dispersive wave tail behind the steep bore-like leading face. Thus despite an apparently different forcing mechanism it appears that a shoaling ISW, or indeed a travelling hydraulic jump at the seabed, may display the properties of a solibore. We thus extend the use of the term throughout this paper to refer to near-bed features that exhibit bore-like features in conjunction with short-period internal waves.

In this paper we highlight two prominent occasions where the occurrence of solibores are evident as a strong temperature front at the seabed and which are followed by, on the first occasion, a turbulent intrusion of cold, dense water with wave-like features on the density interface and, on the second occasion, a wave train whose structure is consistent with a dispersive wave tail. The period separating the two events is ~4 days, and in conjunction with the large scale currents, suggests that they are not forced by the tide despite being in a region where an internal tide is known to exist (Sherwin, 1991). We seek to establish to what extent the solibores are consistent with the results above for shoaling ISWs or overturning hydraulic jumps given their proximity to the seabed, and their likely forcing mechanism. Section 2 outlines the study region and the instrumentation deployed before the results from the field study are presented in Section 3. In Section 4 the cause of the solibores and their implications for near-bed mixing and transport processes over slopes are discussed before conclusions are made in Section 5.

2. Study region and instrumentation

2.1. Study region

Observations were conducted in the Faeroe-Shetland Channel (FSC) during the multi-disciplinary project, Processes over the Continental Slope (PROCS). The aim of PROCS was to study mixing processes over the continental slope, particularly those due to internal wave interactions with the sloping topography, and the influence these have on the cross-slope zonation of sediment

and benthic fauna. A cruise, PROCS 99-1, was conducted in 1999 during the period 14 April–5 May. A variety of moorings were deployed during the cruise and during a subsequent 'long term' 4-month period (PROCS 99-2). Given the predominantly cross-slope nature of the processes to be studied, particularly the internal tide, the moorings were deployed along a transect perpendicular to the continental slope on the southern side of the FSC (Fig. 1) between depths of 471–1000 m (Fig. 2). Given the prominence of the slope in influencing the physical regime, we adopt the standard convention of redefining our axes in the long-slope and cross-slope directions. Positive long-slope (y), referred to throughout the paper as a poleward long-slope flow (V), is defined as a heading of 52°N (True North), whilst positive cross-slope flow (x), referred to as up-slope flow (U), is defined as a heading of 142°N. The vertical coordinate is taken opposite to the direction of gravity, vertically upward from the bottom ($z = 0$ m).

2.2. Instrumentation

Table 1 summarises the details of the instrumentation used and the location of each mooring. Both traditional rotor and vane-type (Aanderaa RCM-8) and acoustic current meters (Aanderaa RCM-9) were mounted on the short-bottom moorings (c3, c5, c6, c8) at $z = 8$, 21, 34 and 47 m. The mechanical current meters exhibited short-period reductions in current speed which have been shown to be most likely caused by turbulence in the near-bed region (Hosegood and van Haren, 2003), and it is the unusually fast sampling of once per minute that provided such clear evidence of the solibores studied in this paper. Optical backscatter sensors (OBS) and sediment traps also measured the light transmittance and sediment fluxes at $z = 2$ and 30 m. A long ranger 75 kHz ADCP (LR-ADCP) was deployed at intermediate positions over the slope (b4), sampling current speed throughout the lower

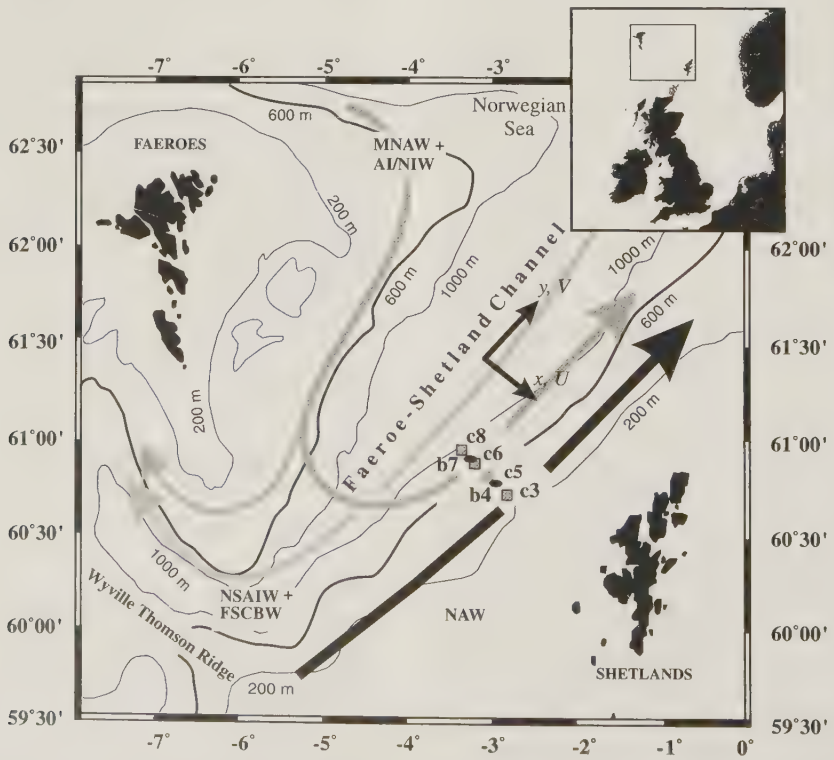


Fig. 1. Map of study region, indicating position of moorings over the slope and the principal water masses that flow through the FSC (see Turrell et al. (1999) for details on hydrography of FSC).

~300 m of the water column. High resolution measurements of current velocity and temperature were enabled by the use of a 600 kHz ADCP and NIOZ thermistor string (van Haren et al., 2001) at 494 m depth (a2).

In addition to moored in situ measurements, two CTD, OBS and Fast Light YoYo (FLY II) free-falling microstructure probe (Dewey et al., 1987) transects were conducted across the channel. At each station the FLY cast followed the CTD cast by ~30–60 min, depending on the water depth. Thus whilst the time interval between casts is relatively short, they should not be considered simultaneous, in particular given the timescale of the phenomenon studied here. We refer briefly to the results obtained from the long-term moorings. The majority of these moorings were lost due to fishing activity so that data is limited to the LR-ADCP at 850 m and sediment traps at 800 and 1050 m.

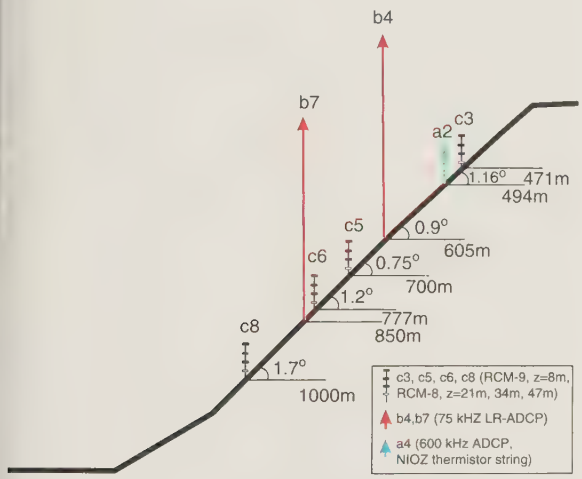


Fig. 2. Schematic of mooring locations over the slope.

3. Observations

In this section we present observations of the two solibores observed travelling up the Shetland slope during PROCS 99-1 on days 112 and 116 and to which we refer throughout the paper as the S1 and S2, respectively. Consideration is given to a third signal (S3) in the currents only at the deeper moorings and which is suggestive of a similar event on day 108. Due to the lack of a definitive temperature signal, however, we do not dwell in detail on S3.

3.1. Temperature and current variability over the slope

The leading edge of solibores S1 and S2 on days 112 and 116 are observed as strong temperature

Table 1
Summary of instrumentation employed during PROCS 99-1(c3, c5, c6, c8, a2, b4) and PROCS 99-2 (d3, d5, b4), between days 108–120 and 120–260, respectively

Mooring and position	Instrument	Depth (m)	Height above bottom, z (m)	Sampling interval (s)
c3	RCM-9, -8, -8, -8	471	8, 21, 34, 47	60
60°49′N, 2°59′W	OBS, PPS $\frac{3}{4}$ trap		2, 30	240, 1 day
a2	RDI 600 kHz ADCP	494	3–30, 0.5 m bins	30
60°49′N, 3°00′W	NIOZ thermistor string[a]		2–33, 1 m intervals	30
b4	RDI 75 kHz ADCP	605	17–357	300
60°53′N, 3°06′W			4 m bins	
d3, 5 (long-term mooring)	PPS $\frac{3}{4}$ trap	800, 1050	2, 30	11 days
60°57′N, 3°14′W; 61°00′N, 3°19′W				
b7 (long-term mooring)	RDI 75 kHz ADCP	850	17–357	600
60°58′N, 3°15′W			4 m bins	

All Aanderaa RCMs were equipped with temperature sensors.
As c3: c5, 700 m, 60°55′N, 3°11′W; c6, 777 m, 60°57′N, 3°13′W; c8, 1000 m, 61°00±N, 3°18′W.
CTD: SeaBird 911.
Microstructure: FLY II free falling probe.
[a]van Haren et al. (2001).

fronts marking the leading edge of the solibores at 471 m, with drops in temperature at $z = 8$ m of 4.47 °C in one minute and 0.9 °C in 10 min (decreasing by 1.674 °C in 51 min) for S1 and S2, respectively (Fig. 3A). There is no temperature signal attributable to either S1 or S2 at the deeper moorings due to their position below the thermocline where the temperature is comparatively uniform in the range $-1\,°C < T < 0\,°C$. Over the ~4 day period prior to day 112, the temperature at $z = 8$ m at 471 m is observed to rise almost monotonically from ~7 °C to a value of 8.377 °C before the drop attributed to S1. Following S1, the temperature again rises during a ~4 day period until S2 to 7.461 °C, when it again drops suddenly with the passage of a solibore.

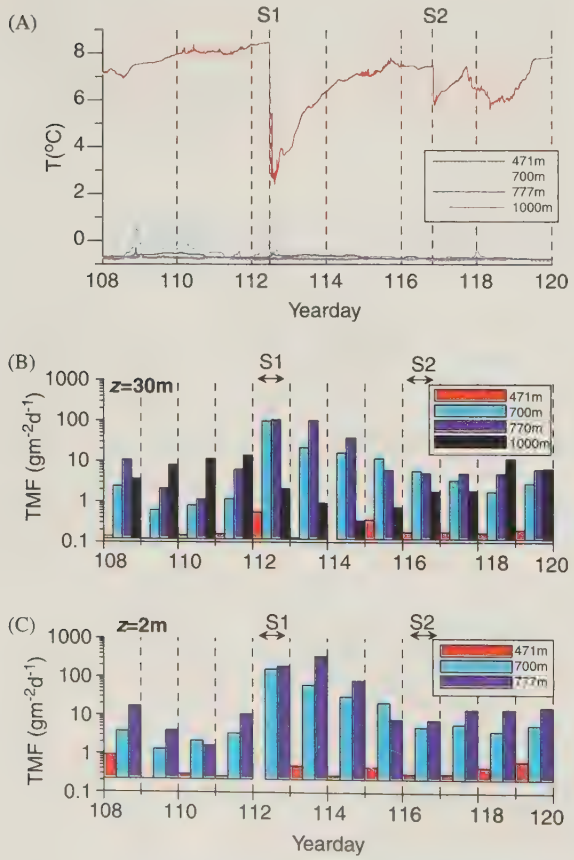


Fig. 3. (A) Temperature from $z = 8$ m, and sediment fluxes from (B) $z = 30$ m and, (C) $z = 2$ m, at 471, 700, 777 and 1000 m. S1 and S2 refer to the times of arrival of the solibores at 471 m.

The sediment fluxes at $z = 30$ m (Fig. 3B) indicate the coincident increase in resuspension associated with the passage of S1, particularly at 700 and 777 m. The largest observed total mass flux (TMF) of 350 g m^{-2} d^{-1} occurs at $z = 2$ m at 777 m on day 113 and is $O(10^2)$ larger than the background values. The time lag between S1 and the maximum sediment flux at this position is presumably due to the settling out of sediment brought into suspension, but the large fluxes at $z = 30$ m indicate the energy and particle motion associated with the solibore was sufficient to enable its upward mixing rather than simple resuspension and settling in the near-bed region.

The passage of S1 in a predominantly upslope direction is implied by the particle velocities measured at $z = 8$ m at 471 m (Fig. 4A); as it passes through the mooring, the cross-slope velocity reverses within ~10 min from a down-slope orientation with an amplitude of > 20 cm s^{-1} to an up-slope direction with a maximum amplitude of 52 cm s^{-1}. During S2 the cross-slope velocity is intensified in the up-slope direction, increasing from a velocity of 10 cm s^{-1} to a peak magnitude of 31 cm s^{-1} within a similarly short-time scale to S1. The only noticeable signal in the cross-slope velocity component at the deeper moorings is the strong reversal at 777 m during S1 from a down-slope flow of 43 cm s^{-1} to an up-slope flow of 35 cm s^{-1} (Fig. 4C), albeit over a slightly longer timescale of ~60 min compared to that at 471 m. A sharp but relatively small drop in the down-slope velocity at 700 m (Fig. 4B) occurs at S1, but is not so pronounced as the signals at either 471 m or 777 m.

The long-slope flow (Figs. 4E and F) is noticeable for the distinct reversal during S1 from an equatorward flow to a poleward flow at 700 and 777 m. The mean flow below the pycnocline is equatorward, so that the reversal during S1 represents a departure from the typical flow. At all depths other than 1000 m the passage of S1 is marked by a poleward surge, whereby the long-slope velocity component increases in strength and persists for a considerable length of time relative to the time-scale of the solibore itself. At each three locations the poleward surge is maintained over a period exceeding 18 h and for nearly 3 days at 471

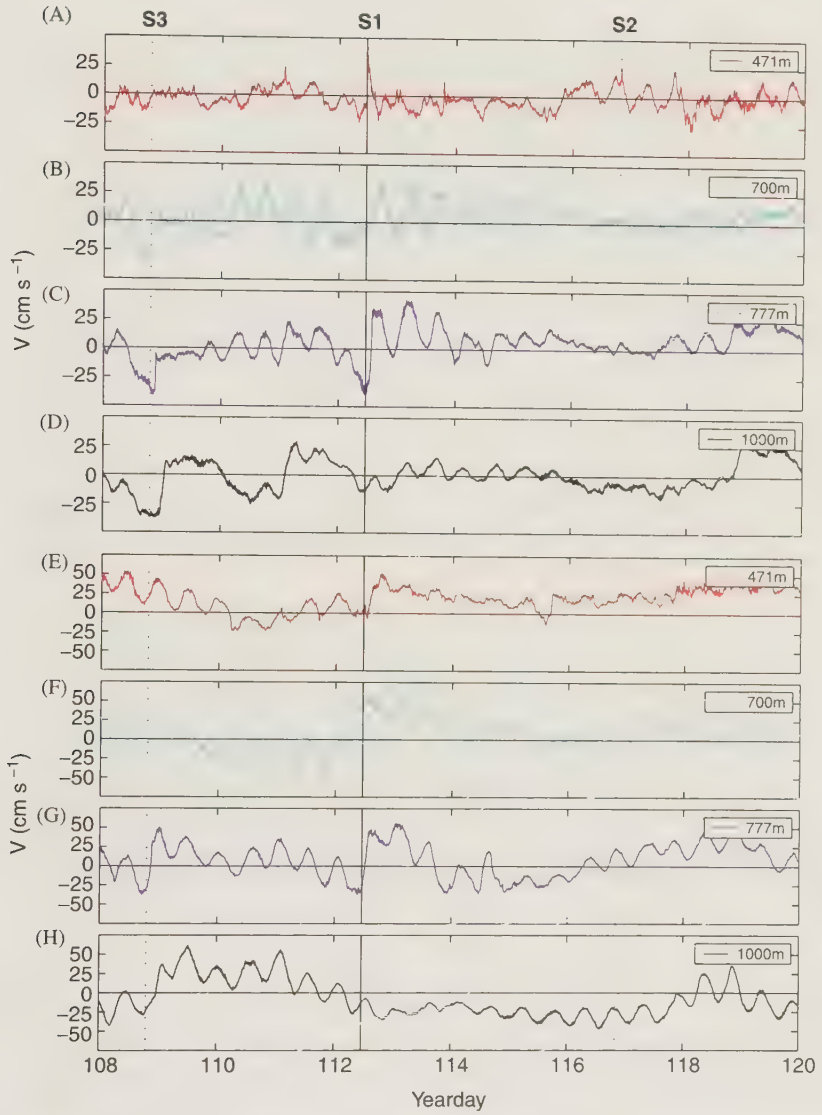


Fig. 4. (A–D) Cross-slope velocity and, (E–H) long-slope velocity from $z = 8$ at 471, 700, 777 and 1000 m. Data are smoothed with a 5-point running mean filter.

and 700 m. Fig. 5 illustrates the subinertial flow at $z = 8$ m at all moorings. The poleward surge following S1 is clear at 471, 700 and 777 m but not 1000 m. The long-slope velocities again turn poleward at the time of S2, although more gradually than during S1. The largest variance is observed at the mid-depth moorings at 700 and 777 m where the flow direction is influenced by the position and depth of the thermocline. It is again clear that 1000 m is out of phase with the shallower moorings.

We have previously suggested that a solibore may have also appeared over the slope on day 108 (S3) despite the lack of any distinct signal in the temperature records at 471 m where the solibores are most apparent due to the strong temperature gradients. We expect that, at the time of S3, the thermocline was constrained at a position over the

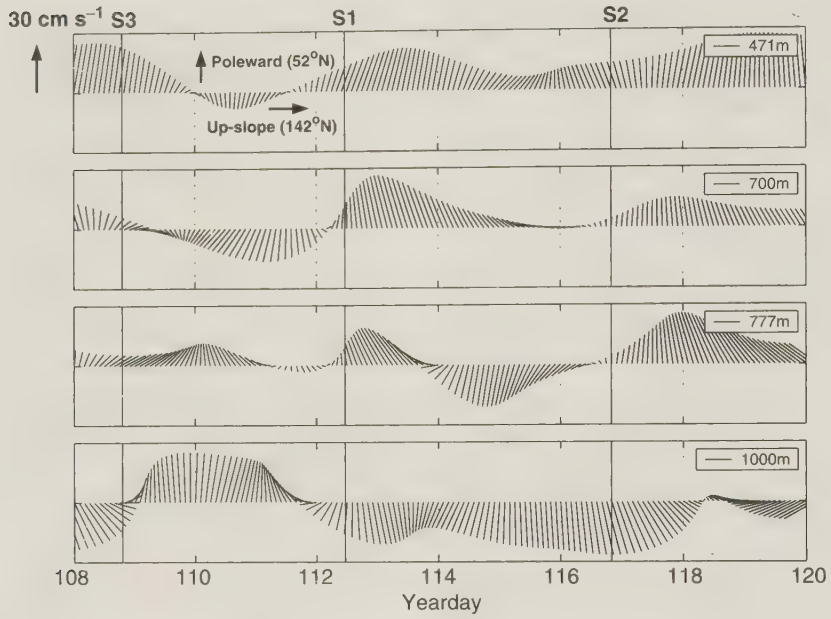


Fig. 5. Low-pass filtered (0.7 cycles/day cut-off) velocity vectors, instruments as in Fig. 4. Vectors are plotted every 2 h.

slope between a2 (494 m depth) and c5 (700 m depth), and that as a result there were simply no temperature sensors available to sample the well-defined temperature gradients so clearly observed during S1 and S2. Sharp gradients in both velocity components at 700, 777 and 1000 m during day 108 are similar to those at 777 m during S1, with the down-slope component at each three locations exhibiting a sudden drop in amplitude of $\sim 30\,\mathrm{cm\,s^{-1}}$ and the poleward surge evident as the long-slope component at 777 m again reverses rapidly from an equatorward flow to a poleward orientation with amplitudes in each direction $> 25\,\mathrm{cm\,s^{-1}}$.

3.2. Current variability with depth at mid-slope

An indication of the mid-slope barotropic currents is obtained by depth averaging the currents over the LR-ADCP range (Fig. 6), although by not reaching the surface the estimate should not be taken as a true value of the barotropic signal. Although S1 is evident as a period of strong currents near the sea-bed ($z < 50$ m) (Figs. 6A and C), it occurs during a barotropic neap tide (Figs. 6B and D), with S2 occurring as spring tide is approached. Near-bed currents at the sea-bed during S2 and S3 are intensified in the long-slope component but not to the same extent as during S1. The core of up-slope, poleward velocities during S1 is trapped within $z < 50$ m, extending upwards to a height of 150 m during day 113 before the signal becomes lost in the background field (Figs. 6A and C). The peak long-slope velocity of $45\,\mathrm{cm\,s^{-1}}$ during S1 leads that at 471 m by approximately 5.3 h whilst the aforementioned poleward surge following S1 is again clear at 605 m on day 112.31.

3.3. Temperature, turbulence and turbidity over the slope

We were fortunate to be conducting a CTD and FLY transect at the time of S1 and were thus able to capture, to a large extent, its passage up the slope and the associated enhanced turbulence it generated (Fig. 7A). A second transect was conducted 3 days later just before S2 and indicates conditions considered more typical of the region (Fig. 7B), but unfortunately prohibits a direct

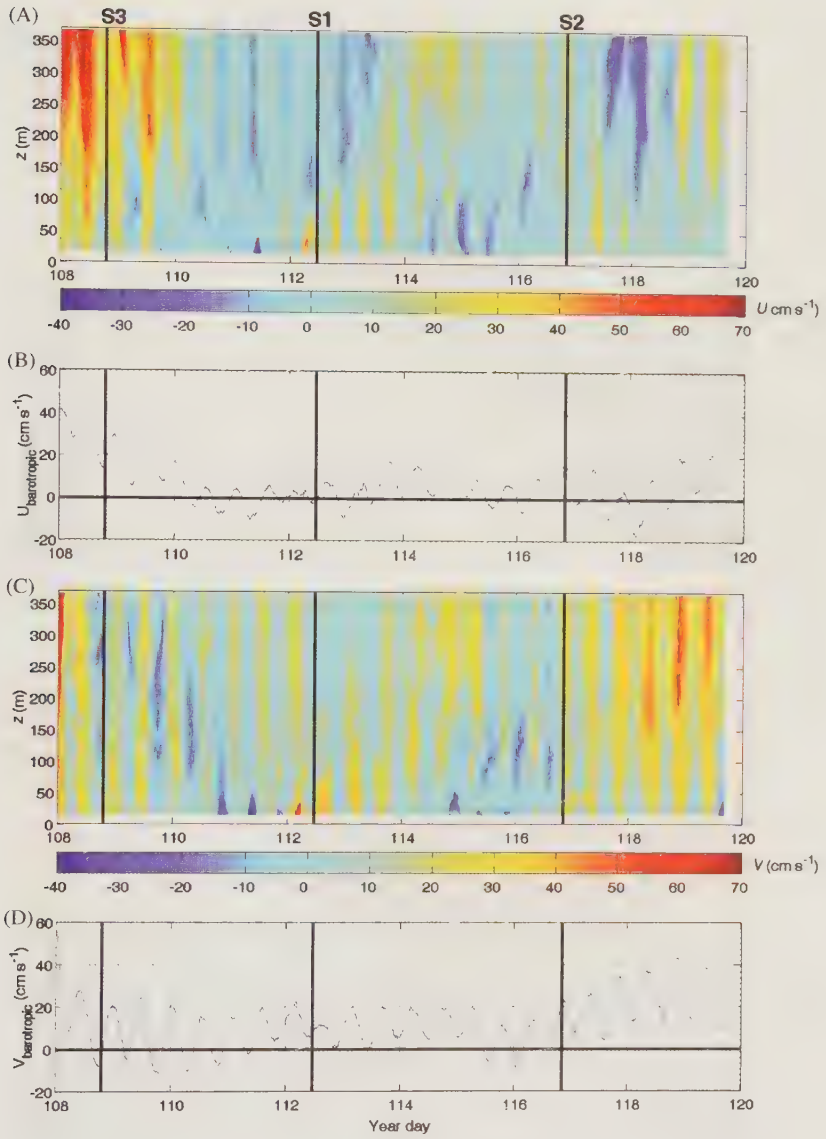


Fig. 6. (A) Cross-slope and (C) long-slope velocity from 75 kHz ADCP at 605 m, and (B) depth-averaged (approximation to barotropic currents) U and (D) depth-averaged V.

comparison of the temperature fields and associated turbulence between S1 and S2 due to the second transect being conducted before, rather than during, the passage of the solibore as was the case for S1. S1 is obvious as an intrusive thin layer (extending to $z = 50$ m) of cold water ($<4\,^{\circ}$C) extending up the slope. The intrusion extends from a depth of ~600 m where the thermocline approaches the slope before running parallel with, and finally intersecting, the slope at a depth of ~500 m. Rates of dissipation of turbulent kinetic energy, ε, typically exceed $10^{-7}\,\mathrm{W\,kg^{-1}}$ in the vicinity of the intrusion. Near the bed they exceed $0.5 \times 10^{-6}\,\mathrm{W\,kg^{-1}}$. The region of enhanced turbulence extends from the sea-bed to the upper edge of the thermocline where the strongest stratification

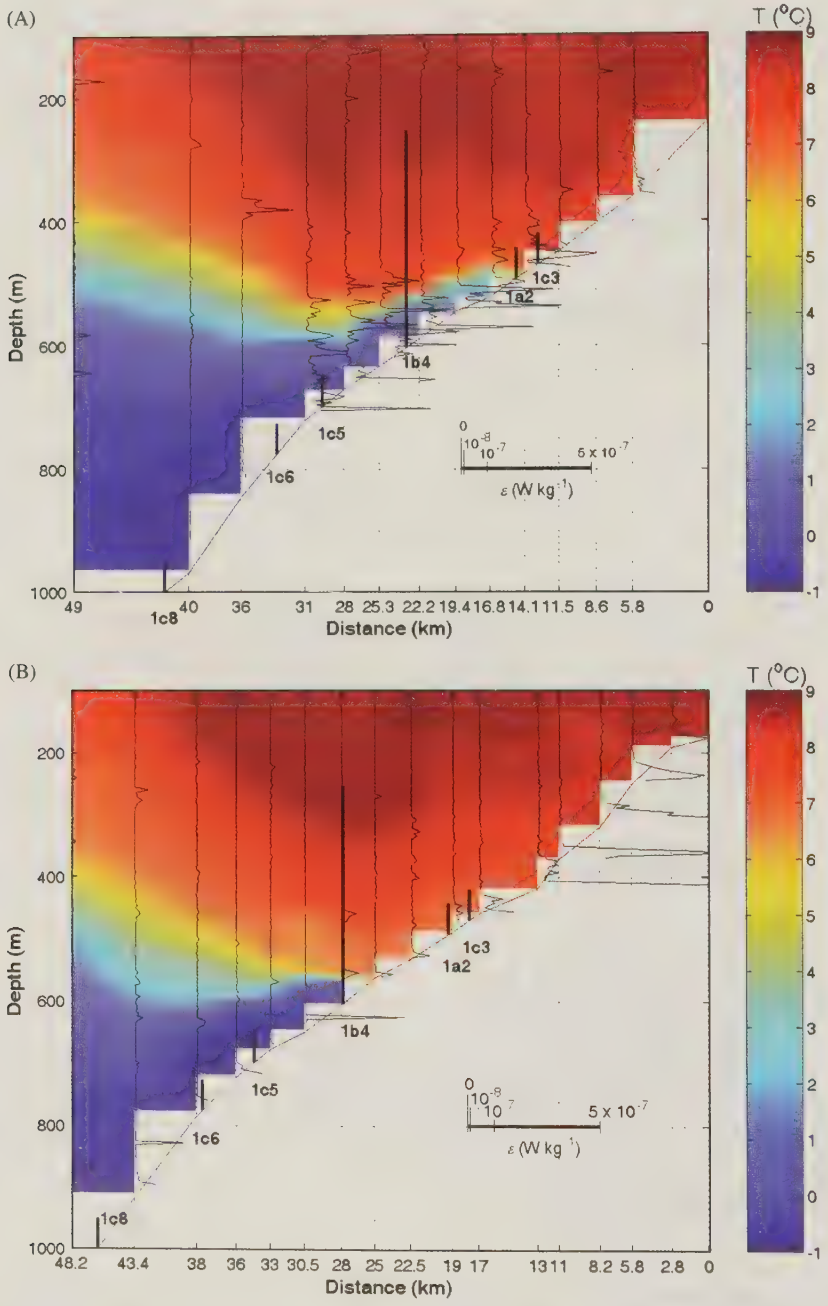


Fig. 7. Temperature and kinetic energy dissipation rate profiles conducted during PROCS 99-1; (A) Transect 1, days 111.97–112.68, and (B) Transect 2, days 116.19–117.06. Profiles were conducted at the specified distances, taken arbitrarily from the shallowest CTD profile.

enhances shear and thus potentially shear instabilities. The weaker stratification near the bed and high values of ε facilitates large vertical eddy diffusivities, $K_z = 0.2\varepsilon/N^2$ (Osborn, 1980). A maximum of $K_z \approx 10^{-1}\,\mathrm{m^2\,s^{-1}}$ is observed at $z = 10\,\mathrm{m}$ at 620 m, whilst K_z at $z < 30\,\mathrm{m}$ is consistently

$>10^{-4}\,m^2\,s^{-1}$. In contrast, the second transect indicates a relatively horizontal thermocline impinging on the slope at a depth of ~550 m and accompanied by low levels of $\varepsilon \approx O(10^{-9})\,W\,kg^{-1}$ (the noise level of this FLY) in the near-bed region within the vicinity of the thermocline. A zone of enhanced resuspension can be clearly seen at a depth of 600–700 m during the first transect (Fig. 8), correlating well with the sediment trap data indicated in Fig. 3.

3.4. High-resolution observations of two different solibores, S1 and S2

In this section we focus on the high-resolution measurements enabled by the use of the thermistor string and ADCP on mooring A2 at 494 m depth. We treat the two distinct solibores separately, beginning with S1 on day 112, before describing the different form of S2 on day 116.

3.4.1. S1

The strong temperature front associated with the leading edge of S1 is seen to extend to over $z = 32$ m at 494 m (Fig. 9A). The largest temperature gradient, $dT/dt_{max} = -2.24\,°C\,min^{-1}$ is observed at $z = 2.5$ m, whilst at $z = 32$ m, dT/dt_{max} is reduced to $-0.699\,°C\,min^{-1}$ at the leading face. The initial front is followed by a packet of approximately 5 waves supported by the strong stratification as indicated for reference by the 6 °C isotherm in Figs. 9B–D and which signifies the thermocline separating the cold (<5 °C) water typical of the intermediate waters in the FSC from the warmer (~7–8 °C) surface waters. The waves are asymmetric and extend upwards into the thicker (450 m) overlying layer. Amplitudes (measured from the preceding trough to the maximum displacement of the 6 °C isotherm) $\eta = O(10\,m)$, the largest being wave 5 with $\eta = 19$ m (Table 2). Periods are $O(30$ min), with corresponding lengthscales of 174–319 m assuming a constant phase speed, c, of $24\,cm\,s^{-1}$, determined by the difference in times of arrival of the temperature front at 494 and 471 m of 81 min. This estimate of phase speed is only valid on the assumption that the particle velocities in the wave front are in the direction of propagation, as is the case for ISW, and that the solibore is thus travelling directly up the slope. Taking this assumption to be valid, the calculation of the phase speed is able to be determined accurately

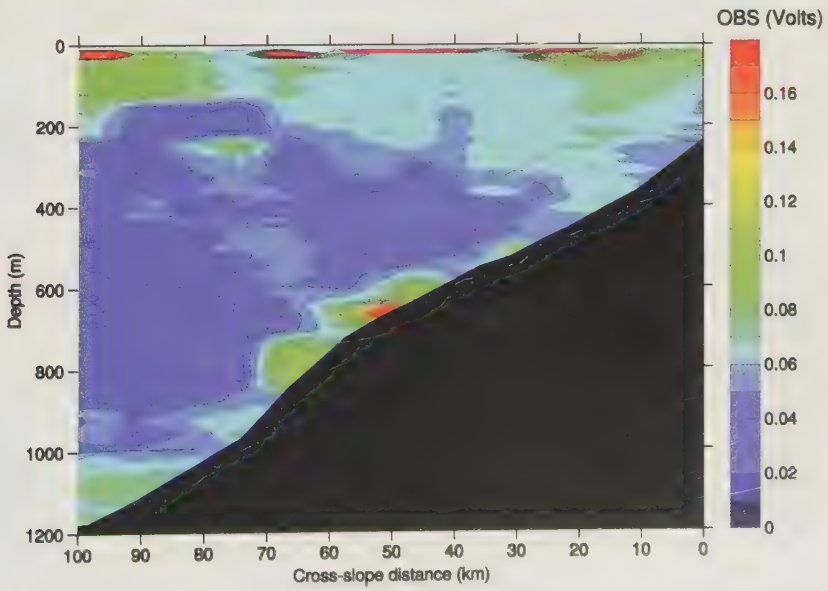


Fig. 8. Light attenuation, measured by the OBS mounted on the CTD frame, during Transect 1 and indicating the core of turbidity associated with S1.

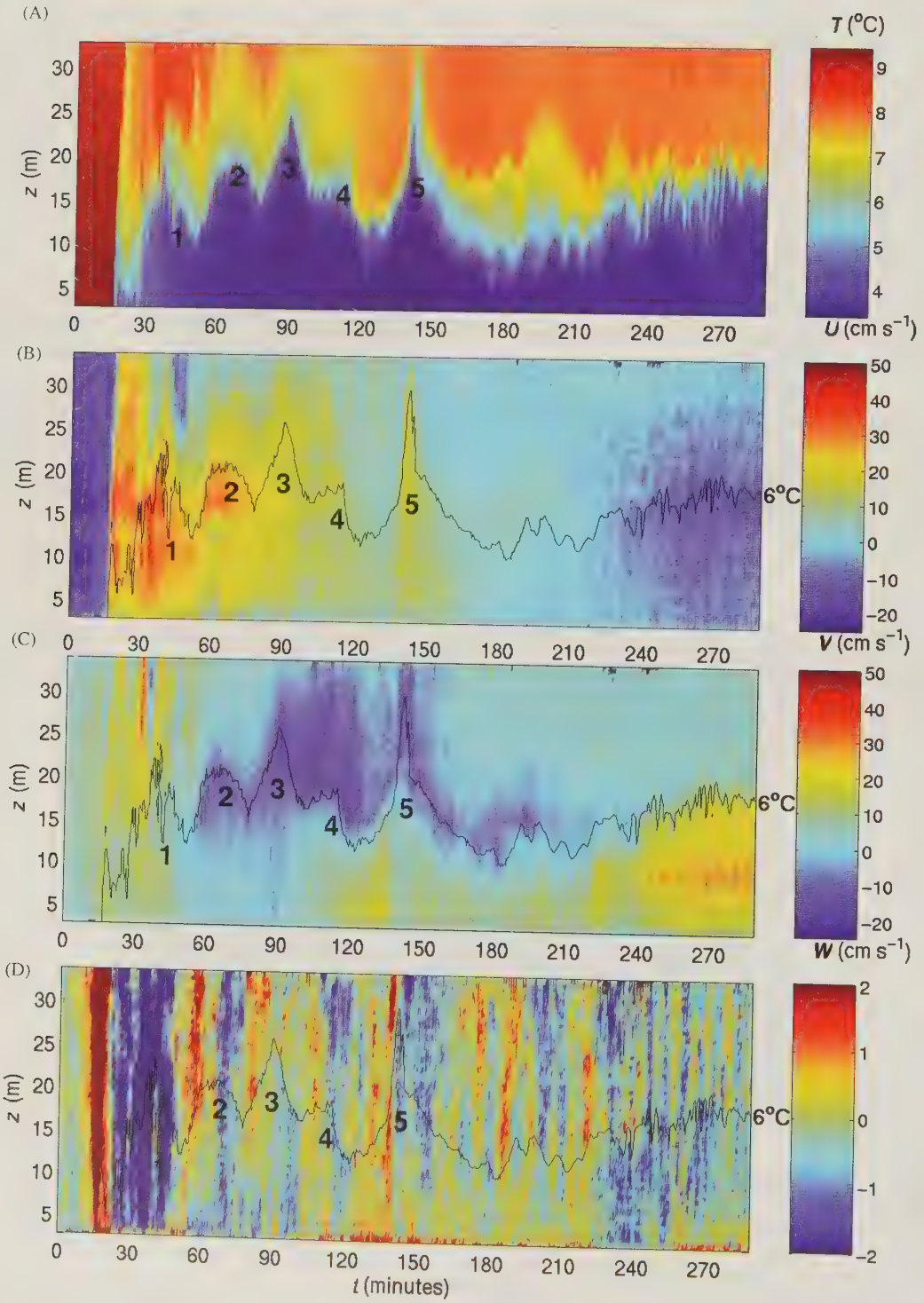

(A)
T (°C)
z (m)
U (cm s−1)
(B)
6°C
V (cm s−1)
(C)
W (cm s−1)
(D)
t (minutes)
1
2
3
4
5

Table 2
Wave characteristics, S1

Wave	Amplitude, η (m)	Wavelength, L, at 0.42η (m)	Period, T (min) at 0.42η
1	18.2	319	20
2	9.2	301	20
3	10.5	174	12
4	7.9	217	14
5	19.1	143	10

The amplitude, η, is defined as the maximum vertical excursion of the 6 °C isotherm from the position of the preceeding trough. Horizontal wavelength is defined as the distance between points on the leading and trailing edges of the wave where the 6 °C isotherm is displaced by $0.42 \times \eta$ (e.g. Osborn and Burch, 1980). Wavelengths are calculated on the premise of a phase speed, $c = 0.25\,\mathrm{cm\,s^{-1}}$.

due to the strength of the temperature gradients being such that the time of arrival of the solibore is able to measured to within a minute corresponding to the instruments sampling period. A further property of an internal bore is that the linear long-wave phase speed, c_{linear}, behind the bore exceeds the speed of the bore itself (Henyey and Hoering, 1997). An appropriate CTD profile, taken at 485 m during the passage of S1 and which indicates an approximately two layer system of two uniform layers separated by a strong density interface, is used to calculate the upper and lower layer densities and thicknesses, ρ_1, H_1 and ρ_2, H_2, respectively, to determine c_{linear}

$$c_{\mathrm{linear}} = \left[\frac{g' H_1 H_2}{(H_1 + H_2)}\right]^{1/2}, \quad (1)$$

where $g' = g(\rho_2 - \rho_1)/\rho_2$. We find for $\rho_1 = 1027.43\,\mathrm{kg\,m^{-3}}$, $H_1 = 450\,\mathrm{m}$, and $\rho_2 = 1027.95\,\mathrm{kg\,m^{-3}}$, $H_2 = 100\,\mathrm{m}$, $c_{\mathrm{linear}} = 0.64 \pm 18\,\mathrm{cm\,s^{-1}}$; a thinner lower layer reduces c_{linear}. Thus S1 appears to satisfy the definition of an internal bore.

The velocity field is characterised by a core of enhanced up-slope velocity at $15 < t < 90$ min (Fig. 9B) reaching a maximum of $U = 52\,\mathrm{cm\,s^{-1}}$ at $z = 16\,\mathrm{m}$ at the leading edge. The front is characterised by a zone of strong convergence, with a down-slope flow immediately preceding the solibore. Superimposed on the core of enhanced velocities are the wave-induced particle velocities, which are most clearly observed during the passage of waves 2 and 5 as perturbations in the up-slope direction. During the passage of wave 5 the cross-slope velocity increases from approximately 0 to $>20\,\mathrm{cm\,s^{-1}}$ at $z = 15\,\mathrm{m}$. The long-slope velocities are comparatively low (Fig. 9C) implying a principally up-slope propagation of S1 if assuming particles velocities are in accordance with those expected for ISWs and further suggested by the time lag between the arrival of the temperature front at 494 and 471 m. The surge in long-slope velocity emerges in the bottom layer at $t = 100$ min, extending upwards away from the bed with time. In contrast long-slope velocity above the interface is directed equatorward. Strong upward vertical velocities are evident in the convergence zone at the leading edge of S1 at $t = 15$ min and extend beyond $z = 32\,\mathrm{m}$ (Fig. 9D). A maximum value of $W = 16.2\,\mathrm{cm\,s^{-1}}$ is observed at $z = 16\,\mathrm{m}$ and is followed by strong downward velocities which persist until $t = 55$ min; thus the leading edge forms a rotor such as proposed by Thorpe (1998) to occur for shoaling internal waves in a lake and whose flow is contrary to the forward breaking of a surface wave. A series of oscillations may then be discerned with typical velocities of $O(2\,\mathrm{cm\,s^{-1}})$ and which correlate with the position of isotherm displacement associated with the waves; upwards (positive) vertical velocities are found at the leading edge of the waves and downwards (negative) velocities at the trailing edge. Note that the vertical velocity oscillations extend to the lowest height above the seabed we were able to measure at $z = 2\,\mathrm{m}$.

Fig. 9. (A) Temperature, (B) cross-slope, (C) long-slope and (D) vertical velocities during S1 as measured at pro1a2 (494 m). Temperature is measured by the NIOZ thermistor string (van Haren et al., 2001), $z = 2$–32 m, velocity by the 600 kHz ADCP mounted at $z = 2$ m. The time origin is taken arbitrarily just before the initial temperature front. Numbers corresponding to individual waves are referred to in the text. The colour scale in (D) is compressed within the maximum and minimum values of W in order to highlight the oscillations in the wave train following the initial rotor.

The density at 494 m is calculated by establishing a temperature–salinity relationship from the CTD profile conducted during the passage of S1 at a depth of 524 m. The station was chosen to ensure the T–S relationship is suitable for the range of temperatures sampled during the passage of the solibores, approximately $3 < T < 8\,^{\circ}\mathrm{C}$ (Fig. 10). A fifth-order polynomial is fitted to the CTD data and the coefficients used to evaluate the density from the thermistor string data and from which the buoyancy frequency is calculated

$$N = \left(-\frac{g}{\rho_0} \frac{\Delta\rho}{\Delta z} \right)^{1/2}, \tag{2}$$

where g is the gravitational acceleration, ρ_0 is a reference density and $\Delta\rho$ the change in density over a depth interval, $\Delta z = 5$ m. A maximum of

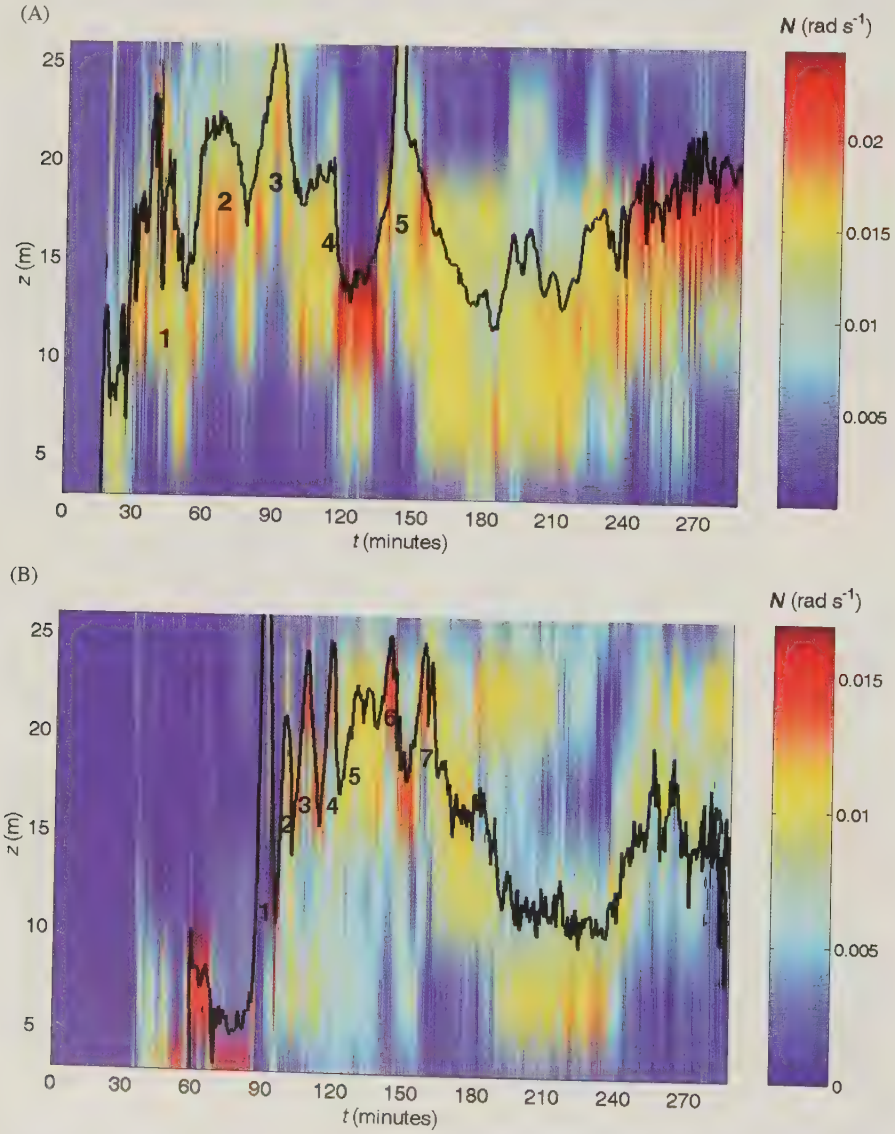


Fig. 10. Buoyancy frequency, N (rad s^{-1}) during (A) S1 and (B) S2. Note the different colour scale used; N was calculated over a vertical distance, dz, of 5 m for each one minute time step.

$N = 2.5 \times 10^{-2}\,\mathrm{rad\,s^{-1}}$ is observed in the trough preceding wave 5, with typical values of $O(10^{-2}\,\mathrm{rad\,s^{-1}})$ found along the thermocline which support the waves (Fig. 10A). At $t = 270$ min a region of strong stratification, with $N_{\max} = 2.4 \times 10^{-2}\,\mathrm{rad\,s^{-1}}$, coincides with a packet of short-period oscillations, following which the depth of the thermocline rapidly shallows and inundates the thermistor string with the colder water. The corresponding buoyancy oscillation period is 260 s whilst the typical period of the high-frequency waves is 100 s; thus the waves have frequencies, $\sigma > N$, which is theoretically impossible given the allowable frequency range of internal gravity waves of $f < \sigma < N$. Two possibilities explain the apparently high frequency of these waves; firstly, the waves may be Doppler shifted to a higher frequency, in a similar manner to that suggested by Thorpe (1987), by the solibore itself due to it being a moving source. Secondly, the waves may be indicative of Kelvin–Helmholtz instabilities resulting from the shear across the density interface rather than being free internal waves. The latter is supported by the slightly tilted orientation of the waves with respect to the vertical, indicative of static instability, and, particularly in the case of the wave-like disturbances at $t = 180-210$ min and $z = 10$ m, 'claw' structures in the temperature record indicative of the formation of vortices at the interface. Thus the apparent waves towards the end of the solibore may actually be billows resulting from the enhanced shear.

3.4.2. High-resolution velocity and temperature measurements, S2

The waves during S2 are shorter but more distinct than S1, with a packet of approximately 7 led by a particularly large wave of amplitude $\eta = 30$ m (Fig. 10B, Table 3). The temperature gradients are weaker given $6\,^{\circ}\mathrm{C} < T < 8\,^{\circ}C$, with $\mathrm{d}T/\mathrm{d}t_{\max} = -0.785\,^{\circ}\mathrm{C\,min^{-1}}$ at $z = 31$ m on the leading edge of wave 1. N is subsequently smaller, reaching a maximum of $1.69 \times 10^{-2}\,\mathrm{rad\,s^{-1}}$ (Fig. 10B). The wave train, defined by the displacement of the 7 °C isotherm, is more coherent than during S1, with waves 1–4 in particular displaying the sech^2 profile typical of ISWs (Fig. 11A). A short extension of strongly stratified ($> 1.6 \times 10^{-2}\,\mathrm{rad\,s^{-1}}$) fluid constrained at $z < 10$ m leads wave 1 at $t = 60-90$ min. The 'nose' is noticeable in the ADCP data as a sharp perturbation in both the cross-slope and long-slope velocities directed up-slope and poleward at $t = 30$ min, as opposed to the ensuing wave train which has perturbations principally in the long-slope component (Fig. 11B). Furthermore, and in contrast to S1, the wave train that trails the initial front is characterised by a core of enhanced velocity in a predominantly poleward direction (Fig. 11C), reaching $V = 40\,\mathrm{cm\,s^{-1}}$ between $120 < t < 180$ min, and with perturbations also in the poleward direction as opposed to the principally cross-slope direction of the wave-induced perturbations in S1. Waves 1–4 all exhibit particle velocities in accordance with ISW (see, for example, Osborn and Burch, 1980) of elevation propagating in an oblique but principally poleward direction; vertical velocities, W, are well defined although, as with S1, the magnitudes are relatively small of $O(2\,\mathrm{cm\,s^{-1}})$ following the leading front which forms a rotor in the same way as S1 (Fig. 11D). Maximum vertical velocities are nonetheless high, with $W = 14.5\,\mathrm{cm\,s^{-1}}$ observed at $z = 27$ m and a return downward flow of $-12.7\,\mathrm{cm\,s^{-1}}$ at $z = 22.5$ m. As opposed to S1 the vertical velocity oscillations are constrained to the vertical extent of the waves and do not extend below the interface to the sea-bed.

Table 3
As Table 2 for S2

Wave	Amplitude, η (m)	Period, T (min) at 0.42η
1	30	4
2	10.5	4
3	10.4	4
4	9.2	4
5	5.5	5
6	4.8	5
7	7.2	4

Wavelength is not given due to our inability to accurately determine the phase speed, c.

3.5. Long-term observations

We briefly summarise the pertinent results from the deep long-term moorings deployed for ~4

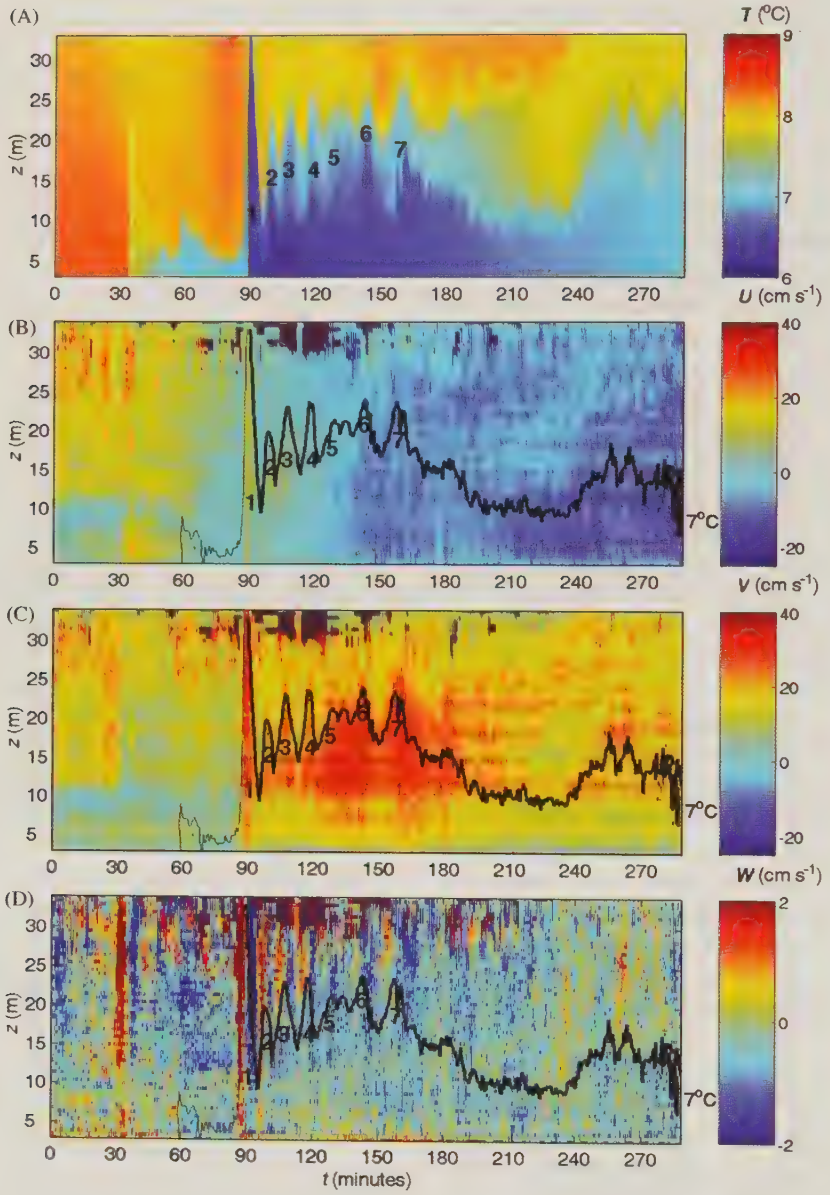


Fig. 11. As Fig. 9 during S2.

months between the two cruises. Spectra computed for each 4 m depth bin of the LR-ADCP deployed at 850 m indicate a concentration of sub-inertial energy between $100 < z < 200$ m at $\sigma \sim 0.35 \pm 0.04$ cpd, corresponding to a period of 2.84 ± 0.35 days (Fig. 12). The concentration of low-frequency energy around a depth of 700 m was also alluded to in Section 3.1 where largest variations were observed at 700 and 777 m. At all frequencies energy decreases towards the bed. The low-pass filtered (cut-off frequency = 0.7 cpd) velocity further demonstrates the existence of the sub-inertial variability in the bottom 300 m of the water column and the sharp gradients in both components that characterised the occurrence of S1 and S2 in the short-term moorings. The

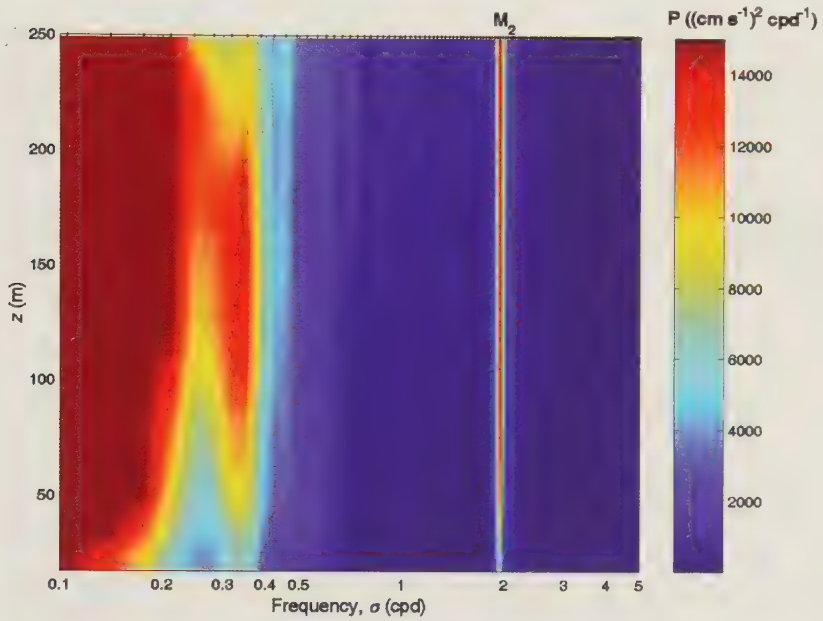


Fig. 12. Horizontal kinetic energy spectra (degrees of freedom = 17) for each depth bin (dz = 4 m) during long-term deployment, days 120–260, from 75 kHz ADCP at 850 m. Tick marks on the top y-axis represent the frequency resolution = 0.0352 cycles/day.

subinertial currents are not intensified at the bottom and there is no phase propagation in the strong signals, which appear to be related to the generation of solibores in the same way as S1 and S2 (Figs. 13A and B). Nonetheless, a period of hugely enhanced resuspension occurs during the period 153.5–164.5 when a flux of 94.14 g m^{-2} d^{-1} (averaged over 11 days) was observed at $z = 2$ m at 800 m (Fig. 13C). During this time the long-slope current reverses rapidly from a poleward to an equatorward direction. The change is as abrupt as during S1 but in the opposite sense. A second period of enhanced resuspension occurs at 1000 m during the period 197.5–219.5 when observed fluxes were 20.03 and 22.44 g m^{-2} d^{-1}; again the long-slope velocity exhibits an abrupt reversal but on this occasion in the same sense as S1, reversing from an equatorward to a poleward direction.

4. Discussion

On two occasions, S1 and S2, within a 12-day period we have observed thermocline intrusions upward along the sloping bottom of the FSC, accompanied by nonlinear waves. We describe them as solibores, implying they begin borelike before becoming more wavelike later in their evolution. Within this section we discuss the potential origin of the solibores and the reason for the differences between the two, their potential for influencing boundary mixing processes and sediment resuspension and redistribution at the continental slope. Finally we propose their potential forcing mechanisms.

4.1. Internal hydraulic jumps at the sea-bed

Holloway (1987) observed hydraulic jumps on the Australian North West Shelf resulting from the steepening of the leading face of an internal tide as it approached shallow water. The isotherms prior to a jump are bunched together indicating a strong stratification above a deep mixed layer, with strong flow directed offshore in the upper layer before rapidly swinging round to an onshore direction with the passing of the hydraulic jump. In our observations, the depression of the thermocline

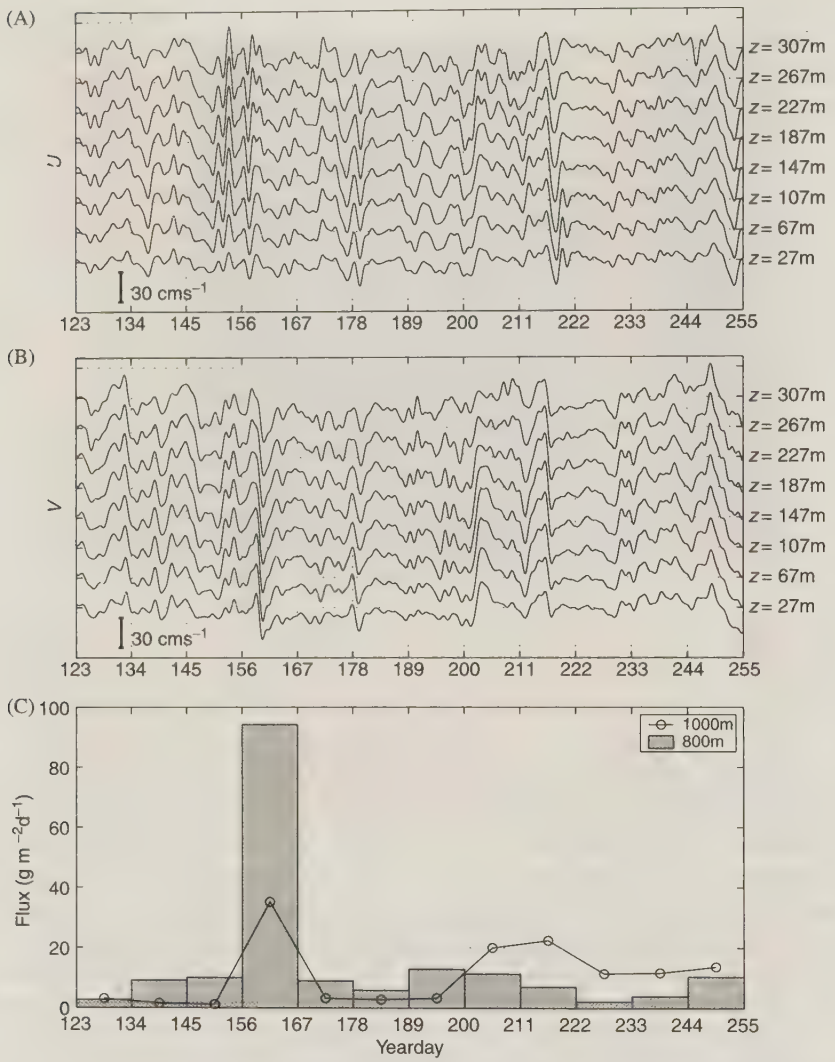


Fig. 13. Every 10th depth bin of low-pass filtered (cut-off frequency = 0.7 cycles/day) (A) cross-slope and (B) long-slope velocity sampled by 75 kHz ADCP at 850 m depth and (C) total mass fluxes measured by sediment traps at $z = 2$ m (800 and 1000 m), during long-term deployment. The velocity for each height is offset by $+30\,\mathrm{cm\,s^{-1}}$ from the previous record.

prior to S1 implies a strong, near-bed stratification, whilst on day 111.4 we observe a distinct pulse in down-slope, equatorward flow (Figs. 14C and D) near the bed at 605 m and which is immediately followed by a reversal in current direction to an upslope, poleward orientation. Fig. 6 indicates that the pulse is concentrated at the bottom and thus, we assume, represents flow in the lower layer below the pycnocline. We consider whether the pulse is sufficiently strong to cause a hydraulic jump.

Following Holloway (1987) but reversing the density interface to a near-bed position as opposed to a near-surface pycnocline, we define the nondimensional variables (see Fig. 15A)

$$\text{Layer height}: \quad R_1 = h_1/H; \quad R_2 = h_2/H,$$
$$\text{Velocity ratio}: \quad T_1^2 = u_1^2/u_2^2; \quad T_2^2 = u_2^2/u_1^2,$$

where h_1, h_2 are the upper and lower layer thicknesses, H is the total water depth and u_1, u_2

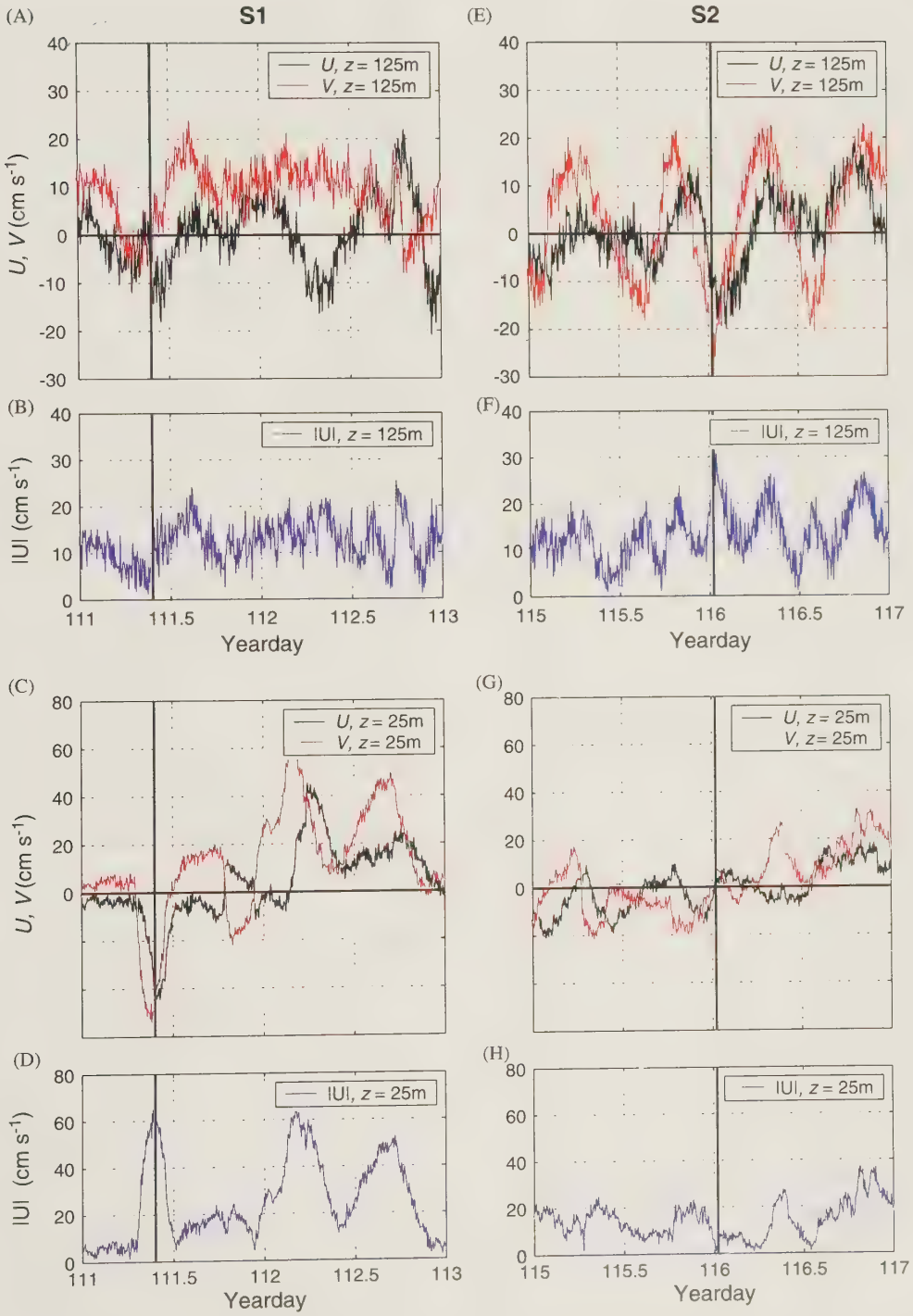


Fig. 14. (A) Cross-slope, U, and long-slope, V, velocity, and velocity amplitude, $|U|$ during S1 (A–D) and S2 (E–H) at $z = 25$ m and 125 m, 605 m depth. Solid vertical lines at year days 111.4 and 116.02 correspond to the times for which calculations of hydraulic conditions were made.

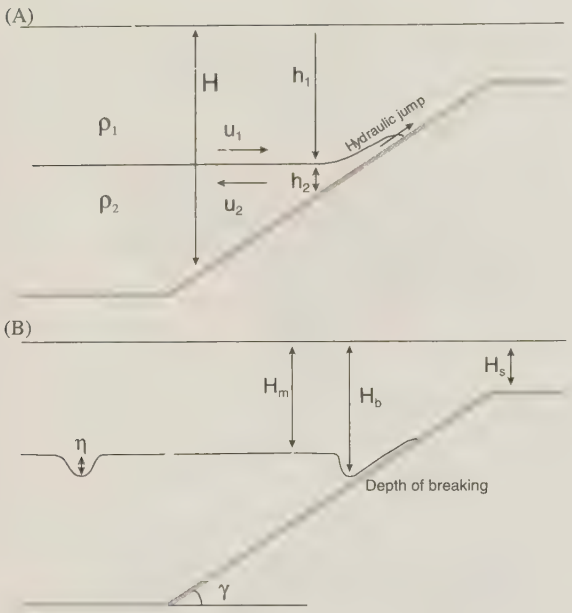


Fig. 15. Schematic of parameters used in: (A) hydraulic model and (B) model for shoaling ISWs.

are the upper and lower velocities, respectively. The Froude number, Fr, may be expressed as

$$Fr = \left[1 + \frac{RT^2}{1 - R^2}\right]^{-1/2}, \quad (3)$$

for which a jump is expected for hydraulically critical conditions when $Fr \leqslant 1$. Applying the observed values of u_1, $u_2 = 20$, $60\,\mathrm{cm\,s^{-1}}$, respectively, $h_2 = 50\,\mathrm{m}$, $H = 650\,\mathrm{m}$ and g' as calculated for (1), gives $Fr = 0.99$ applying R_2 and H_2 to give the lower layer Froude number. The low value of $R = 0.07$ ensures that $Fr \sim 1$ for a large parameter space, such that even for variations in T^2 hydraulically critical conditions are likely. Thus the pulse in down-slope, equatorward flow prior to S1 on day 111.4 seemed sufficient to create a hydraulic jump. Prior to S2 the temperature data across the slope imply the thermocline was not so depressed, but a pulse is observed at $z = 125\,\mathrm{m}$ with relatively quiescent conditions at $z = 25\,\mathrm{m}$. Applying R_1 and H_1 to give the upper layer Froude number because the strong flow is above the thermocline, $Fr = 0.89$ on assuming the same values of h_1 and h_2, again suggesting hydraulically critical conditions were achieved and promoting the development of a hydraulic jump. The orientation of the jump is not dependent on which layer is hydraulically critical, as confirmed by Holloway (1987) who found both upward and downward jumps for critical conditions in the upper layer alone.

Whilst the evolution of the internal tide across the Australian North West shelf as a hydraulic jump to an undular bore has been studied (Smyth and Holloway, 1988), the proximity of the hydraulic jump to the sea-bed means that a more suitable context within which to discuss our findings is the evolution of a baroclinic bore at the sea-bed, studied in laboratory and numerical experiments on shoaling ISWs over slopes. Vlasenko and Hutter (2002) (hereafter VH02) found that the *rear* face of an incident ISW of amplitude η steepens due to nonlinearity (Fig. 15B), forming a baroclinic bore which may then overturn due to kinematic instabilities. The overturned bore then transforms into a horizontal density intrusion propagating up the slope. For less steep slopes, however, the longer transition zone enables dispersion to balance nonlinearity and form a dispersive wave tail. A number of similarities exist between our observations over a sloping boundary in the ocean and the numerical results of VH02. The similarities we observe are

(1) A rotor is formed at the leading edge of the bore, with strong upward velocities followed by a region of downward velocities of slightly smaller magnitude (Figs. 9D, 11D; Figs. 3G–I in VH02).
(2) After breaking the ISW transforms into a horizontal density intrusion propagating onto the shelf, with a core of strong horizontal velocity in the direction of propagation (Figs. 9B, 11C; Fig. 7B in VH02).
(3) Wave-like structures appear along the upper edge of the density interface bounding the horizontal density intrusion, and are the result of the multiple reflections from the bottom following overturning (Fig. 9B; Fig. 7A in VH02). Such structures are also referred to as 'soliton-like waves of elevation' (VH02) or

turbulent boluses (Helfrich, 1992; Lamb, 2002).

(4) The effect of the multiple reflections from the sea-bed is a pattern of oscillating vertical velocities which extend to the sea-bed (Fig. 9D; Fig. 7C in VH02) rather than being constrained to the leading and trailing edges of genuine ISWs.

(5) For a gentle slope a dispersive wave tail, as is observed in S2, is formed rather than an overturning hydraulic jump (Fig. 11; Fig. 9 in VH02).

The parameters that describe the breaking event in the numerical studies of VH02 are unlikely to be obtained in the ocean with the required accuracy to make quantitative comparisons with the theory useful. This is the case here, where a lack of measurements in the far field and of high resolution measurements deeper than those obtained at 494 m and where the hydraulic jump is expected to have originated, render a determination of the ISW amplitude required for breaking to occur impossible. Furthermore, the aforementioned pulse in current speed prior to S1 in particular and the subsequent strong reversal in direction is not consistent with a shoaling ISW. However, the results of VH02 essentially relate to an overturning baroclinic bore and may be extended to the case of hydraulic jumps at the sea-bed resulting from other processes.

The importance of the bottom slope to the breaking process described in VH02 is useful to elucidate qualitatively the difference between the forms of S1 and S2. An ISW propagating from deep water into a region of sloping topography will break at the depth H_b (Fig. 15B), when the relationship

$$|a| = \frac{\eta}{H_b - H_m}, \tag{4}$$

where H_m is the undisturbed depth of the isopycnal exhibiting maximum displacement by the ISW and $|a|$ is the nondimensional wave amplitude, is satisfied by

$$|a| \cong \frac{0.8^\circ}{\gamma} + 0.4, \tag{5}$$

where γ is the angle of the bottom-slope (VH02). Thus as γ is reduced in (5), the magnitude of $|a|$ is increased, such that a larger initial wave amplitude is required for breaking to occur. The direction of propagation of S1 inferred from the particle velocities is principally up-slope along a bottom having an observed $\gamma = 1.16^\circ$, whilst the particle velocities of $>26\,\mathrm{cm\,s^{-1}}$ are greater than the speed of the bore which we calculate by comparing the time of arrival of the temperature front at 494 m and then at 471 m. This satisfies the kinematic instability mechanism for the overturning of a baroclinic bore according to Kao et al. (1985), Grue et al. (2000) and VH02 and explains the form of S1 being consistent with the results of VH02 for an overturned ISW (points 1–4 above). During S2, however, when the form of the solibore is more that of a dispersive wave train, the particle velocities imply a principally long-slope direction of propagation; at $t = 90$ min we observe $U = 9\,\mathrm{cm\,s^{-1}}$ and $V = 32\,\mathrm{cm\,s^{-1}}$, giving an angle between isobaths and the boundary projection of the group velocity vector, $\theta = 16^\circ$ (Fig. 16). The effective slope is thus reduced to $\gamma_{\mathrm{effective}} = \sin^{-1}(\sin\gamma \sin\theta) = 0.3^\circ$, which for shoaling ISWs increases the amplitude required for breaking to occur but, from the perspective of nonlinearly steepening bores and hydraulic jumps, allows dispersion time to balance the nonlinearity and form a dispersive wave tail (point 5 above). The observed value of $\gamma_{\mathrm{effective}}$ is furthermore very similar to the angle of 0.52° for which a dispersive wave tail is found to evolve in the numerical results of VH02. It is the influence of the reduced effective

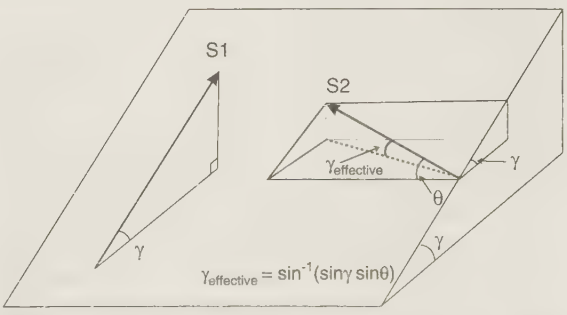


Fig. 16. Schematic indicating the reduced effective slope felt by the solibore during S2 due to an oblique angle of propagation, θ.

slope on the overturning process through the competition between nonlinearity and dispersion that we propose to be the essential difference in the form of S1 and S2.

4.2. Implications for near-bed mixing and sediment resuspension

Transect 1 provided insight into the degree of mixing resulting from an overturning hydraulic jump. Values of $K_z = 10^{-1}\,m^2\,s^{-1}$ are observed at $z = 10\,m$, with more typical values of $10^{-4}\,m^2\,s^{-1}$ sustained throughout the 4 h duration of the solibore. Such a value is of the order of magnitude proposed by Munk and Wunsch (1998) to be required to maintain the thermohaline circulation; however as has been discussed throughout the paper, the solibores occur with a typical period of 4 days and a duration of the order of hours. Thus they are unlikely to contribute significantly to the globally averaged diapycnal mixing. They do, however, have important consequences for mixing in lakes, where wind-forced seiches may occur on a regular basis and degenerate into packets of ISWs (Horn et al., 2001; Boegman et al., 2003) which on reaching the sloping boundary will almost certainly break due to a steeper boundary inclination, γ, as compared to the ocean, thus causing localised mixing as they lose up to 70% of their energy on reflection (Michallet and Ivey, 1999). The results presented here indicate the importance of the obliquity of propagation of ISWs with respect to the slope in determining the effective slope over which they shoal. At oblique angles the effective slope is reduced, potentially resulting in the formation of a dispersive wave train rather than an overturning wave.

The consequences for sediment transport in regions of sloping topography appear to be particularly important. Observations from the FSC suggest that the first solibore, S1, facilitated sediment fluxes $O(10^2)$ larger than the background values, implying that such events completely dominate the mid-slope resuspension (see Bonnin et al. (2002) for the details of the sediment analysis). In contrast to the results of Bogucki et al. (1987), who observe resuspension during periods of low bed-shear stress due to near-bed propagating solitons, we observe resuspension to occur during periods of high bed shear stress in a similar way to Cacchione et al. (2002) who propose that continental slopes are effectively shaped by reflecting internal tides which cause near-bed velocities to generate bed shear stresses high enough to inhibit the deposition of fine-grained sediment. Thomsen and Gust (2000) propose a velocity of $45\,cm\,s^{-1}$ at $z = 1\,m$ is required to resuspend sediment proper, whilst at 700 and 777 m we observe velocities of $>70\,cm\,s^{-1}$ at $z = 8\,m$ during S1. However, the currents at other times during the observational period also exceed $45\,cm\,s^{-1}$ but at which time sediment fluxes were *not* observed to increase, requiring a further explanation beyond the bed shear stress mechanism proposed by Cacchione et al. (2002) as to why fluxes during S1 were so high. We propose that two factors are important in causing the large fluxes; firstly, the rotor formed at the leading edge of the bore will not only bring sediment into suspension due to the strong vertical velocities but will further transport it upwards against the density gradient that otherwise inhibits its vertical transport. This appears to be confirmed at 494 m where, despite comparatively minor enhanced fluxes during S1, the echo intensity from the ADCP provides an indication of suspended matter in the water column. Intensity is largest towards the leading edge where the rotor is observed (Fig. 17A), whilst values are comparatively low during S2 during which no enhanced fluxes were observed (Fig. 17B). A similar mechanism is proposed by Johnson et al. (2001) whereby ISW-induced convergent bottom currents induce vertical motions that lift the nepheloid layer away from the bed in shallow water. Secondly, once brought into suspension, the turbulent boluses are proposed by Helfrich (1992) and Lamb (2002) to be potentially very effective transporters of sediment in an upslope direction. Larger particles are expected to be deposited as the boluses decay whilst finer particles settle out over time and may be advected offshore with the mixed fluid. The results of Bonnin (unpublished results) indicate this to be the case, with the sustained fluxes over day 113 (Fig. 3B) composed of finer particles whilst the flux attributed to the larger particles peaks during the

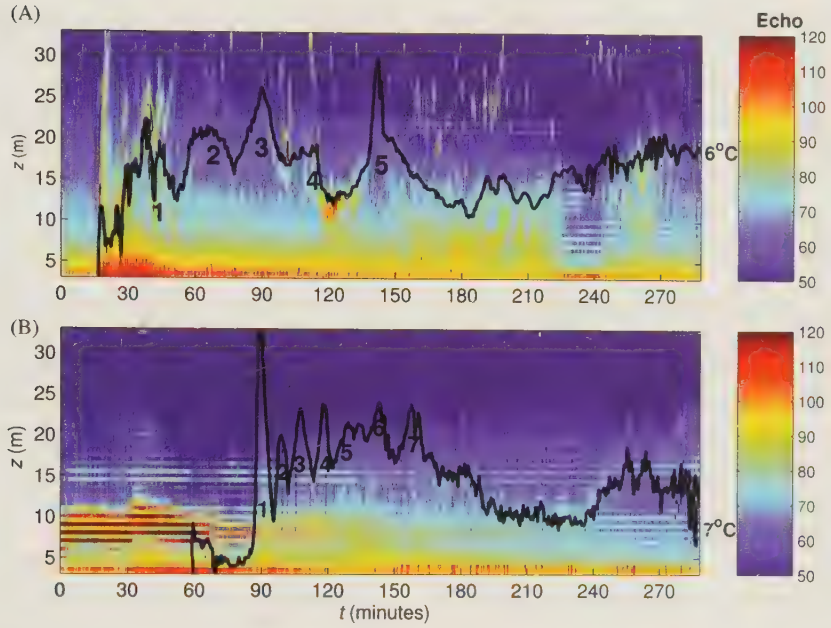


Fig. 17. Echo intensity from ADCP at 494 m during (A) S1 and (B) S2. The mean for each depth bin, $\langle I \rangle$ is removed to account for beam attenuation throughout the water column and values subsequently increased by the minimum value at each depth interval throughout the deployment.

passage of S1 on day 112 and decays immediately following its passage. No enhanced resuspension was observed during S2 when we believe a dispersive wave tail was formed rather than the turbulent boluses. A distinct rotor with comparable vertical velocities was also observed during S2 however, and we propose that the principal reason for the lack of resuspension was the depth at which the solibore originated. The cooler temperature at 471 m depth prior to S2 implies the shallower depression of the thermocline and thus of the point of origin of the solibore as compared to S1; in the FSC available sediment is scoured by the boundary currents at depths of <600 m, with the consequence that there was not any sediment available for resuspension during S2. This further explains why the fluxes during S1 were negligible at 471 m. The long-term (millennial) significance of the up-slope sediment transport has been alluded to by observations of up-slope migrating sedimentary structures in the FSC (Christian Wilson, British Geological Survey, Personal Communication, 2003). High-resolution seismic data indicate a number of agrading mounds over the slope between depths of 500–800 m which are built up on a continuous basis from irregular surfaces, possibly originating from glacial times. Thus the solibore-induced up-slope transport of sediment in the FSC appears to be important to the long-term evolution of the sea-bed topography as well as to the short-term transport of material over the slope.

4.3. *Large scale forcing for solibore generation*

The results presented here are consistent with the evolution of a hydraulic jump propagating up the slope, due to either shoaling ISWs over a sloping bottom or to the apparently hydraulically supercritical flow prior to the jumps. Given the lack of data from further away from the slope to indicate an incident ISW approaching the slope, we thus consider how a hydraulic jump may be formed at the sea-bed. Specifically, we are concerned with a dynamical mechanism that not only forces perturbations in the thermocline or in the near-bed region over the slope, but does so

with a subinertial periodicity of typically 4 days and is accompanied by current surges in the long-slope direction below the thermocline. Similar 5–6 day periodicities were observed by Thorpe (1987) over the continental slope of the Porcupine Bank which appeared to propagate poleward with a phase speed of $13\,\mathrm{cm\,s^{-1}}$; thus such signals appear to be commonplace over the north European continental slope.

Internal tides are discounted despite reports by Sherwin (1991) that a deep internal tide exists in the region and is responsible for periodic 'swirls of sand and dust' as reported by the crew of mini-sub operators in the study region. It was considered whether the depression and intensification of the thermocline leading up to S1 brought the slope region at 700 m to criticality for the M_2 semi-diurnal lunar internal tide. The critical frequency is related to the bottom-slope by (Thorpe, 1987)

$$\frac{\sigma^2}{f^2} = \frac{N^2 \sin^2 \gamma}{f^2} + \cos^2 \gamma, \tag{6}$$

where here σ is the frequency of the M_2 tide. Under critical and near-critical conditions, reflecting internal waves may form internal bores which propagate up the slope (Dauxois and Young, 1999), potentially evolving in a similar way to the process of ISW overturning as the stratification is eroded in a band above the bed, bounded at its outer edge by the internal wave characteristic (Legg and Adcroft, 2003). The problem lies in the timing of S1 during a barotropic neap tide when internal tide generation would be expected to be at a minimum. Gerkema (2002) proposes that the strong baroclinic signal at the sea-bed during a barotropic neap tide may be due to a phase shift in the baroclinic spring-neap cycle with respect to the barotropic cycle. Subsequently, enhanced baroclinic signals may be found at the sea-bed but which are intermittent, sensitive to small changes in the background stratification. This is an attractive idea given the variable periodicity of the solibores in the FSC where the background stratification varies, but we discount it because of the concentration in horizontal kinetic energy at a narrow-band subinertial frequency (O(0.3 cycles/days)) observed at depths removed from the bottom by ~100 m during the long-term deployments (Fig. 12). Furthermore, the CTD profiles indicate the slope at 700 m to be subcritical to the M_2 tide (Fig. 18), despite some isolated spikes in the stratification during the second transect which brought small regions of the slope to criticality. The sustained surges in the long-slope current velocity associated with the solibores are also not consistent with semi-diurnal internal tide generation or reflection. One of the principal problems in determining which process may be responsible for the low frequency forcing is the stratification, and in particular its variability in both space and time. Internal wave characteristics are sensitive to changes in the background density field and which will alter the paths taken by the wave; thus any signal at a given point in space and time due to the internal tide may be intermittent on a time scale unrelated to the tide itself but instead due to the different periodicity of variability in the stratification which dictates whether, in the case of the internal tide, the wave is generated and/or reflected or not.

We have considered the possibility that the bottom intensification of currents is due to a bottom-trapped topographic Rossby wave (Rhines, 1970) and which is also dependent on the strength of stratification. Such waves have been proposed to be responsible for a bottom-trapped, 7-day periodicity in the Gulf of Lions (Flexas et al., 2002) and 2–4 and 4–6 day periodicities in the Santa Barbara Channel (Auad et al., 1998). In the FSC, where γ is typically 1°, and $S = N/f \sim 20$, waves become bottom-trapped at frequencies $\sigma < \gamma S$, corresponding to periods > 2.5 days and with length scales, $L < 20$ km. The FSC is conducive to type 3 waves in Rhines (1970), requiring $\gamma \ll S^{-1}$ and which have $\sigma \leqslant fS^{-1}$ corresponding to a period of ~10 days. As with the case of internal tides above, the frequency is dependent on the stratification and is thus variable. Thus we do not expect such a motion to occupy a specific frequency in the energy spectra but rather to exhibit a broadband response, again at odds with results from the long-term ADCP. Furthermore, the subinertial currents exhibit little decay away from the bottom, except on day 204, and with no phase propagation (Fig. 13). That the currents at

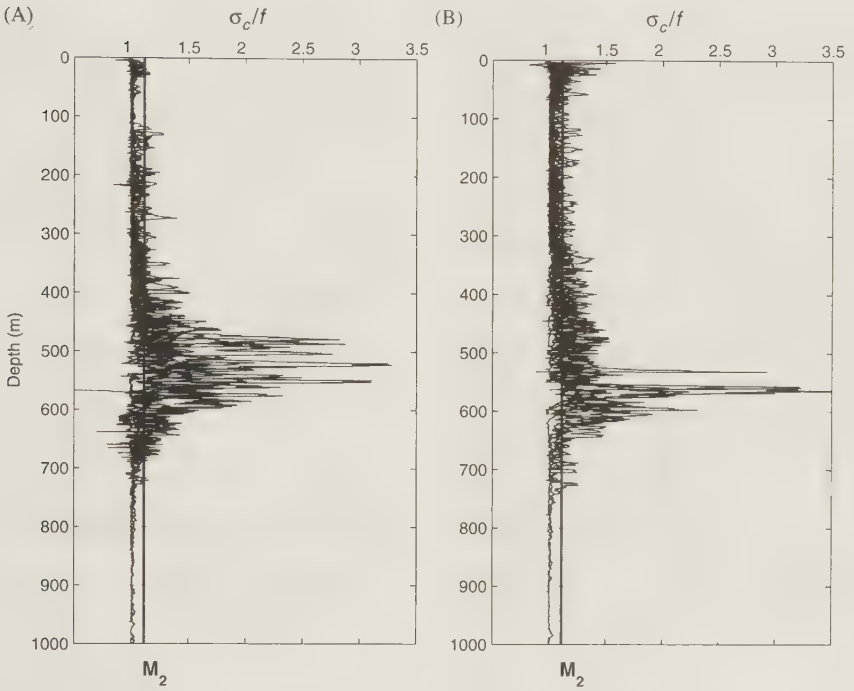


Fig. 18. (A) Critical internal wave frequency, σ_c, normalised over f for CTD profiles conducted during Transect 1 and, (B) Transect 2. Vertical line at $\sigma_c/f = 1.1$ corresponds to the critical frequency for the M_2 internal tide. Profiles are smoothed with a 20 m vertically running mean filter.

605 m are bottom intensified during the passage of the solibores is also subject to debate; the solibores represent the passage of the thermocline up the slope, with the consequence that the flow measured at a particular mooring will be determined by the respective layer within which it is sampling at the time. Thus as the solibore passes a mooring, the instruments will be sampling the lower layer below the thermocline rather than the surface layer prior to the solibore. If the flow in the lower layer is stronger than the upper, the currents measured by an ADCP will appear to be bottom intensified as the lowest bins are in the lower layer, with the signal spreading up as more of the water column is occupied by the fast flowing lower layer.

Thorpe et al. (1995) observe jumps in thermocline depth arising from the interaction of nonlinear internal Kelvin wave fronts with the sloping topography in Lake Geneva which cause a rapid descent of the thermocline and a sudden reversal and intensification of the currents. Along-slope flows may persist for many hours after the wave has passed (Thorpe, 1998). In Section 4.2 it was shown how current pulses prior to S1 and S2 were sufficient to result in hydraulically supercritical conditions, yet in contrast to Holloway (1987) the internal tide is not responsible. We propose that a nonlinear Kelvin wave, propagating to the northeast with the shallow right to the right, may be a potential mechanism facilitating the formation of hydraulic jumps over the sloping bottom, in the same way as proposed by Thorpe (1998). Pressure charts (Fig. 19) indicate that a strong and persistent low-pressure system was situated over the study region from days 110 to112 at which point S1 was observed. As with the observations in Lake Geneva, we expect there to be a lag of approximately 2–3 days, depending on the proximity of the forcing to the study region, between the atmospheric forcing and the oceanic response in the form of the Kelvin wave. Thus it appears that the low pressure system passing to the west of the

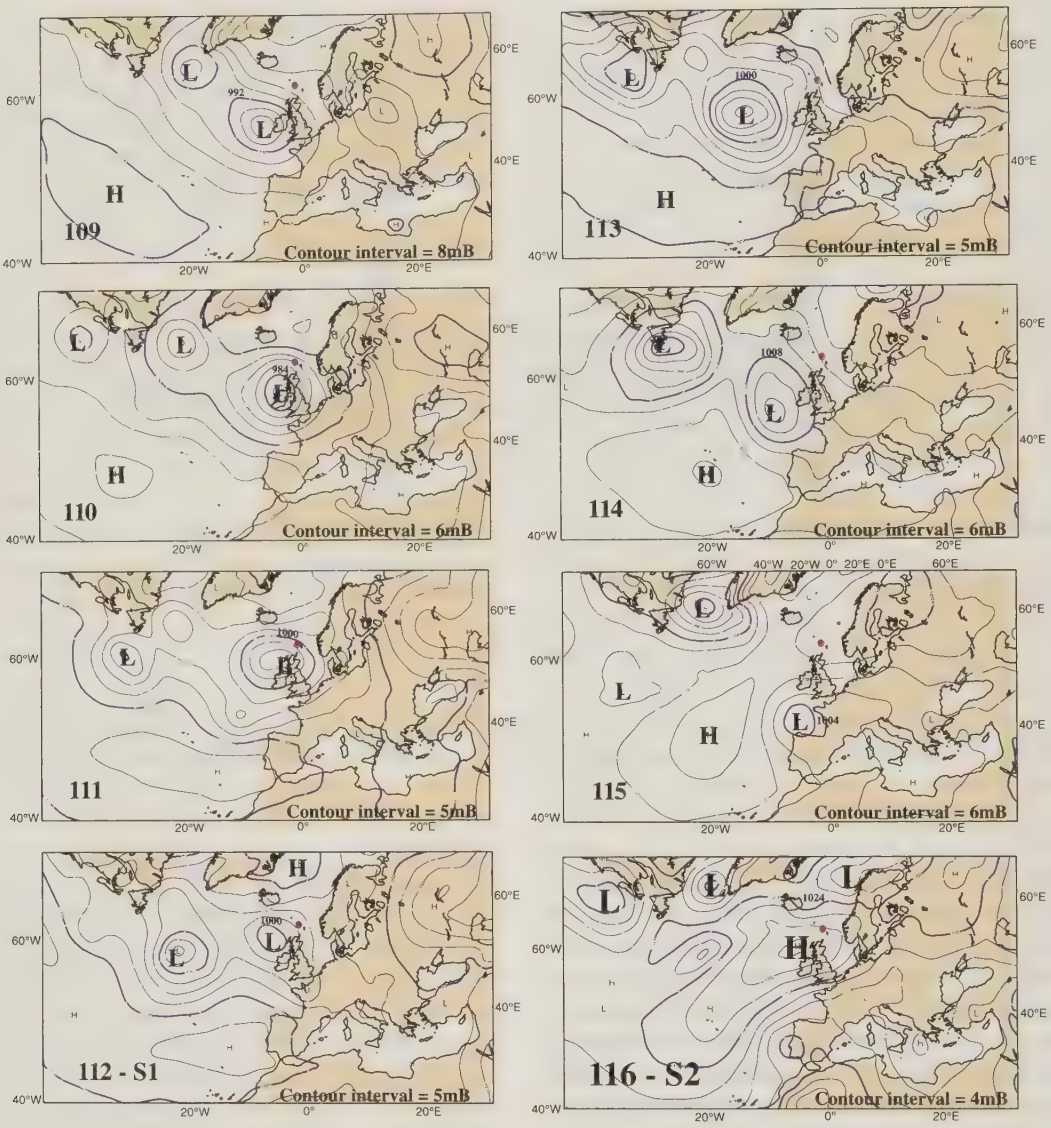


Fig. 19. Mean sea level pressure on days 109–116 (days are indicated in bottom left of each panel), indicating the presence of a persistent low pressure system in the study region prior to S1 up to 112 and a transient low pressure system passing to the west of the study region on days 113 and 114, two days before S2. Low pressure and high pressure are designated by L and H, respectively, and pressure contour intervals are variable; the isobar designated by the thick contour in the study region is labelled and the contour intervals (in mB) are indicated in the bottom right corner of each panel. Approximate position of the moorings are indicated by the red circle. Source: ECMWF ERA-40 Data Set.

study region on days 113 and 114 may be implicated in the forcing of S2. It should be stressed that the different forms of S1 and S2 are due to their response, specifically their angle of propagation with respect to the slope, and not their initial forcing by the Kelvin wave. It is as yet unclear however what factors are important in determining the obliquity of propagation of a solibore resulting from such a process and dedicated future work is needed to elucidate this.

5. Conclusions

During a 12-day measurement period over the continental slope in the FSC, two intense non-linear wave trains were observed propagating up the slope. The wave trains have been described as solibores according to the definition of Henyey and Hoering (1997) with the exception that they do not owe their existence to tidal forcing. They are separated by a period of approximately 4 days and similar signals in the same records and from a long-term deployment suggest a subinertial but variable periodicity to be related to their forcing. We are as yet unable to positively identify the origin of such a periodicity and thus the forcing mechanism of the solibores and propose that the problem lies in an incomplete knowledge of the background stratification to which several of the potential forcing mechanisms are particularly sensitive.

The structure of the solibores is consistent with the results of numerical and laboratory experiments studying overturning hydraulic jumps resulting from shoaling ISWs. Observations of one solibore indicate both the kinematic instability mechanism as being effective and the eventual density and velocity fields to be in accordance with an overturned hydraulic jump. The second solibore propagates at an oblique orientation to the slope however, thereby reducing the effective slope and allowing dispersion to balance the nonlinearity and produce a dispersive wave tail rather than overturning. The density and velocity fields display perturbations with typical solitary wave profiles and particle velocities indicative of genuine solitary waves rather the turbulent boluses resulting from the overturned hydraulic jump. It is shown that hydraulically supercritical conditions were achieved at the sea-bed prior to the solibores, possibly resulting from the passage of a non-linear Kelvin wave, presenting an alternative generation mechanism for the propagating internal hydraulic jump.

The occurrence of such solibores has important implications for understanding the behaviour of internal hydraulic jumps, forced by either shoaling ISWs other sources, propagating over sloping regions in both the oceanic and lacustrine environments. For the latter in particular, ISWs are expected to be significant in the energy transfer from large basin-scale waves to small scales that are available for mixing. In the ocean, the near-bed solibores are very effective at resuspending sediment as shown here by sediment traps deployed near the sea-bed. A rotor, formed by the vertical velocities resulting from the horizontal convergence at the leading edge of the solibore, is proposed to bring sediment into suspension whilst turbulent boluses then transport it up-slope. Support for such a process is provided by seismic data suggesting that the up-slope transport of sediment in the FSC has occurred over a geological timescale and is important in shaping the topography. Given their relative infrequency and short duration, the solibores are not expected to significantly influence the internal wave-induced globally averaged diapycnal diffusivity in the deep-ocean.

Acknowledgments

The data were acquired during the Processes on the Continental Slope (PROCS) project, funded by the Netherlands Organisation for the Advancement of Scientific Research (NWO). We thank the Department of Sea Technology for the preparation of the moorings and the crew of the *RV Pelagia* for their assistance during their deployment and recovery. We are grateful to Jerome Bonnin for the use Figs. 1 and 8 and for his analysis of the sediment traps. We thank Kees Veth for the use of the microstructure data and further thank Theo Hillebrand for his careful and thorough preparation of the instrumentation.

References

Auad, G., Henderschott, M.C., Winant, C.D., 1998. Wind-induced currents and bottom-trapped waves in the Santa Barbara Channel. Journal of Physical Oceanography 28, 85–102.

Boegman, L., Imberger, J., Ivey, G.N., Antenucci, J.P., 2003. High-frequency internal waves in large stratified lakes. Limnology and Oceanography 48 (2), 895–919.

Bogucki, D., Dickey, T., Redekopp, L.G., 1987. Sediment resuspension and mixing by resonantly generated internal

solitary waves. Journal of Physical Oceanography 27, 1181–1196.

Bonnin, J., van Raaphorst, W., Brummer, G-J., van Haren, H., Malschaert, H., 2002. Intense mid-slope resuspension of particulate matter in the Faeroe-Shetland Channel: short-term deployment of near-bottom sediment traps. Deep-Sea Research I 49, 1485–1505.

Bryden, H.L., Candela, J., Kinder, T.H., 1994. Exchange through the strait of Gibraltar. Progress in Oceanography 33, 201–248.

Cacchione, D.A., Pratson, L.F., Ogston, A.S., 2002. The shaping of continental slopes by internal tides. Science 296, 724–727.

Colosi, J.A., Beardsley, R.C., Lynch, J.F., Gawarkiewicz, G., Chjiu, C.-S., Scotti, A., 2001. Observations of nonlinear internal waves on the outer New England continental shelf during the summer Shelfbreak Primer study. Journal of Geophysical Research 106 (C5), 9587–9601.

Dauxois, T., Young, W.R., 1999. Near-critical reflection of internal waves. Journal of Fluid Mechanics 390, 271–295.

Dewey, R.K., Crawford, W.R., Gargett, A.E., Oakey, N.S., 1987. A microstructure instrument for profiling oceanic turbulence in coastal bottom boundary layers. Journal of Atmospheric and Oceanic Technology 4, 288–297.

Djordjevic, V.D., Redekopp, L.G., 1978. The fission and disintegration of internal solitary waves moving over two-dimensional topography. Journal of Physical Oceanography 8, 1016–1024.

Eriksen, C.C., 1982. Observations of internal wave reflection off sloping bottoms. Journal of Geophysical Research 87, 525–538.

Flexas, M.M., Durrieu de Madron, X., Garcia, M.A., Canals, M., Arnau, P., 2002. Flow variability in the Gulf of Lions during the MATER HFF experiment (March–May 1997). Journal of Marine Systems 33–34, 197–214.

Gerkema, T., 2002. Application of an internal-tide generation model to baroclinic spring-neap cycles. Journal of Geophysical Research 107 (C9), 3124.

Grue, J., Jensen, A., Rusas, P.O., Sveen, K.J., 2000. Breaking and broadening of internal solitary waves. Journal of Fluid Mechanics 413, 181–217.

Helfrich, K.R., 1992. Internal solitary wave breaking and run-up on a uniform slope. Journal of Fluid Mechanics 243, 133–154.

Helfrich, K.R., Melville, W.K., Miles, J.W., 1984. On interfacial solitary waves over slowly varying topography. Journal of Fluid Mechanics 149, 305–317.

Henyey, F.S., Hoering, A., 1997. Energetics of borelike internal waves. Journal of Geophysical Research 102 (C2), 3323–3330.

Holloway, P.E., 1987. Internal hydraulic jumps and solitons at the shelf break region on the Australian North West Shelf. Journal of Geophysical Research 92 (C5), 5405–5416.

Horn, D.A., Imberger, J., Ivey, G.N., 2001. The degeneration of large-scale interfacial gravity waves in lakes. Journal of Fluid Mechanics 434, 181–207.

Hosegood, P.J., van Haren, H., 2003. Ekman-induced turbulence over the continental slope in the Faeroe-Shetland Channel as inferred from spikes in current meter observations. Deep-Sea Research I 50 (5), 657 680.

Howell, T.L., Brown, W.S., 1985. Nonlinear internal waves on the California continental shelf. Journal of Geophysical Research 90 (C4), 7256–7264.

Hunkins, K., Fliegel, M., 1973. Internal undular surges in Seneca Lake: a natural occurrence of solitons. Journal of Geophysical Research 78, 539–548.

Jeans, D.R.G., Sherwin, T.J., 2001a. The evolution and energetics of large amplitude nonlinear internal waves on the Portuguese shelf. Journal of Marine Research 59, 327–353.

Jeans, D.R.G., Sherwin, T.J., 2001b. The variability of strongly non-linear solitary internal waves observed during an upwelling season on the Portuguese shelf. Continental Shelf Research 21, 1855 -1878.

Johnson, D., Weidemann, A., Scott Pegau, W., 2001. Internal tidal bores and bottom nepheloid layers. Continental Shelf Research 21, 1473–1484.

Kao, T.W., Pan, F.S., Renouard, D., 1985. Internal solitons on the pycnocline: generation, propagation and shoaling and breaking over a slope. Journal of Fluid Mechanics 159, 19–53.

Lamb, K.G., 2002. A numerical investigation of solitary internal waves with trapped cores formed via shoaling. Journal of Fluid Mechanics 451, 109–144.

Legg, S., Adcroft, A., 2003. Internal wave breaking at concave and convex continental slopes. Journal of Physical Oceanography 33 (11), 2224–2246.

Michallet, H., Ivey, G.N., 1999. Experiments on mixing due to internal solitary waves breaking on uniform slopes. Journal of Geophysical Research 104, 13467–13478.

Morozov, E.G., Trulsen, K., Velarde, M.G., Vlasenko, V.I., 2002. Internal tides in the Strait of Gibraltar. Journal of Physical Oceanography 32, 3193–3206.

Munk, W., 1966. Abyssal recipes. Deep-Sea Research 13, 707–730.

Munk, W., Wunsch, C., 1998. Abyssal recipes II: energetics of tidal and wind mixing. Deep-Sea Research I 45, 1977–2010.

New, A.L., Pingree, R.D., 1990. Evidence for internal tidal mixing near the shelf break in the Bay of Biscay. Deep-Sea Research I 37, 1783–1803.

Osborn, T.R., 1980. Estimates of the local rate of vertical diffusion from dissipation measurements. Journal of Physical Oceanography 10, 83–89.

Osborn, AR., Burch, T.L., 1980. Internal solitons in the Andaman Sea. Science 208 (4443), 451–460.

Pingree, R.D., Mardell, G.T., New, A.L., 1986. Propagation of internal tides from the upper slopes of the Bay of Biscay. Nature 321, 154–158.

Rhines, P., 1970. Edge-, bottom-, and Rossby waves in a rotating stratified fluid. Geophysical Fluid Dynamics 1, 273–302.

Sandstrom, H., Elliot, J.A., 1984. Internal tide and solitons on the Scotion Shelf: a nutrient pump at work. Journal of Geophysical Research 89 (C4), 6415–6426.

Sherwin, T.J., 1991. Evidence of a Deep Internal Tide in the Faeroe-Shetland Channel. In: Parker, B.B. (Ed.), Tidal Hydrodynamics. Wiley, New York, pp. 469–488.

Simpson, J.E., 1997. Gravity Currents in the Environment and the Laboratory, Second ed. Cambridge University Press, Cambridge, UK.

Small, J., Sawyer, T.C., Scott, J.C., 1999. The evolution of an internal bore at the Malin shelf break. Annales Geophysicae 17, 547 565.

Smyth, N.F., Holloway, P.E., 1988. Hydraulic jump and undular bore formation on a shelf break. Journal of Physical Oceanography 18, 947–962.

Thomsen, L., Gust, G., 2000. Sediment erosion thresholds and characteristics of resuspended aggregates on the western European continental margin. Deep-Sea Research I 47 (10), 1881–1897.

Thorpe, S.A., 1977. Turbulence and mixing in a Scottish Loch. Philosophical Transactions of the Royal Society of London A 286, 125–181.

Thorpe, S.A., 1987. Current and temperature variability on the continental slope. Philosophical Transactions of the Royal Society of London A 323, 471–517.

Thorpe, S.A., 1998. Some dynamical effects of internal waves and the sloping sides of lakes. In: Imberger, J. (Ed.), Physical Processes in Lakes and Oceans, American Geophysical Union, Coastal and Estuarine Studies, vol. 54, pp. 441–460.

Thorpe, S.A., Keen, J.M., Jiang, R., Lemmin, U., 1995. High frequency internal waves in Lake Geneva. Philosophical Transactions of the Royal Society of London A 354, 237–257.

Turrell, W.R., Slesser, G., Adams, R.D., Payne, R., Gillibrand, P.A., 1999. Decadal variability in the composition of Faroe Shetland Channel bottom water. Deep-Sea Research I 46, 1–25.

van Haren, H., Groenewegen, R., Laan, M., Koster, B., 2001. A fast and accurate thermistor string. Journal of Atmospheric and Oceanic Technology 18, 256–265.

Vlasenko, V., Hutter, K., 2002. Numerical experiments on the breaking of solitary internal waves over a slope-shelf topography. Journal of Physical Oceanography 32, 1779–1793.

Wallace, B.C., Wilkinson, D.L., 1988. Run-up of internal waves on a gentle slope in a two-layered system. Journal of Fluid Mechanics 191, 419–442.

Wesson, J.C., Gregg, M.C., 1988. Turbulent dissipation in the strait of Gibraltar and associated mixing. In: Nihoul, J.C.J., Jamart, B.M. (Eds.), Small-Scale Turbulence and Mixing in the Ocean. Elsevier, Amsterdam, pp. 201–212.

ELSEVIER

Deep-Sea Research II 51 (2004) 2973–2993

DEEP-SEA RESEARCH
PART II

www.elsevier.com/locate/dsr2

Variability on steep, confined topography

William K. Dewar*, Dmitri A. Leonov

Department of Oceanography, The Florida State University, Tallahassee, FL 32306, USA

Abstract

Analytical f-plane wave characteristics forced by abrupt topography are computed and compared to numerical solutions. The topography is assumed to be both confined and steep and relative vorticity is not neglected, points that contrast the present study with some recent analyses. Both baroclinic and barotropic cases are considered. It is argued that such waves possess a number of unusual linear dispersive and nonlinear steepening tendencies. For example, baroclinic waves on finite topography mix barotropic and baroclinic dynamics in a manner distinct from those on weak topography. Their nonlinear behavior also discriminates strongly between on and offshore directions in amplitude and scale. Westward group velocities for weak topography are typical only for varicose waves, which does seem to characterize some observed waves. Finite topography however, routinely supports westward group velocities for sinuous and varicose waves because it is effective at filtering barotropic dynamics. It is thus suggested that the intersection of isopycnals with topography introduces a qualitative change to the dynamics of topographic variability.
© 2004 Elsevier Ltd. All rights reserved.

1. Introduction

Irregular, rough topography is a prevalent ocean feature, appearing in the shape of the ocean bottom and coastline and in the shelves and slopes that connect the shores to the abyss. Strong currents, jets and ocean variability are also often observed near topography, arguing that bottom bumps, continental shelves and seamounts exert a strong dynamical oceanic influence, but theoretical or quantitative information about such connections is limited. The objectives of this paper are to present a combined analytical and numerical discussion of variability on confined topography, clarify the dynamics behind such variability, illustrate the novel features topography introduces to variability and comment on relevant ocean phenomena. The focus of the analysis is on topography of limited lateral extent (e.g. continental slopes with steep gradients), as opposed to slowly varying topography which has been the subject of some recent developments. In addition, the analysis applies to all wavelengths, not just long wavelengths where relative vorticity is weak or absent. Finally, the analysis is extended into the weakly nonlinear

*Corresponding author.
E-mail address: dewar@ocean.fsu.edu (W.K. Dewar).

0967-0645/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dsr2.2004.09.011

regime and interactions between linear wave dynamics and nonlinear tendencies are examined.

1.1. Background

Past studies of topographic effects break largely into two categories differing in their expression of topographic variability. One camp examines the effects of slowly varying, and hence relatively smooth, topography and the other various types of "rough" topography. The first group of studies often employ long wave assumptions while the latter group includes studies of topographic effects on geostrophic turbulence and the interactions of escarpments and waves. The present paper falls into this last group.

An important precursor study of the present work is Rhines (1969), who examined geostrophic waves on a step discontinuity of depth in a barotropic fluid (see also Longuet-Higgins, 1968). The waves were trapped to the discontinuity, and hence were named "double Kelvin waves". Rhines (1977) discovered a dispersion relation for such waves over a stepped bottom in a two layer fluid, and described the relationship between their vertical structure and wavelength (see also Willmott, 1984). Rhines (1970) examined waves in a continuously stratified system as a function of bottom inclination and argued that Kelvin wave dynamics are often supplanted by topographic wave dynamics, provided the lateral scale of the topography exceeds the deformation radius. Recent contributions to the study of stepped topography include Willmott and Johnson (1989), who considered an application of linearized dynamics to a circular lake with a shelf. Willmott and Johnson (1995) have illustrated the role of the double Kelvin wave in the slow phase of Rossby adjustment over an escarpment.

An interesting oceanic setting for wave-topography interactions was given in Pickart (1995); see also Schultz (1987) who argued fluctuations on the Mid-Atlantic Bight continental slope were caused by Gulf Stream variations northeast of the Bight that propagated pseudo-westward as topographic waves to the observation point. His analysis employed a slowly varying assumption about the topography along the wave trajectory and expanded upon earlier analyses by Hogg (1981) and Pickart and Watts (1990) to show how typical Gulf Stream variability could couple to topographic waves. The value of this result was in offering an explanation for the energetic topographic wave variability of the region, which had been recognized as far back as Thompson and Luyten (1976).

1.2. This paper

Given the strong observational basis supporting topography as a major factor influencing, and possibly controlling, ocean variability, theoretical investigations into topographic impacts on variability find strong motivation. The objective of the present paper is contributed to our analytical understanding of topography and variability. Specifically, connections between the potential vorticity fields associated with abrupt and confined topography and wave evolution are examined. The combined foci of abrupt topography, stratification and nonlinearity are the main ways the present work expands upon past studies. The present methodology also permits the derivation of path equations for the variability. This proves to be a convenient shorthand for discussing wave dynamics.

Accordingly, it is argued that the dominantly observed tendency for topographic variability to propagate energy to the "west" arises either from the finite slope associated with abrupt topography, and is then varicose (i.e., out-of-phase across the slope), or the intersection of isopycnals with the topography, which acts as a barotropic filter. Evidence for varicose structure is found in subsurface current meter observations for some dominant forms of Mid-Atlantic Bight variability, a result consistent with the present analysis. There is also some evidence of westward propagating sinuous variability (i.e., in-phase across the shelf), a result again consistent with the theory and that suggests topography might play a strong role in regional variability. Nonlinearity on finite topography discriminates between onshore and offshore directions. The manner in which this happens is a complicated function of wavenumber, but affects the amplitude and scales of wave

steepening. Numerical calculations support the theory.

Variability on weak topography is discussed in the next section, which also illustrates the analysis technique. The following section considers the case of finite topography involving a generalization of QG to permit isopycnal intersections with topography. The paper ends with a Discussion section, where connections to Mid-Atlantic Bight observations are attempted.

2. Waves on abrupt, but weak, topography

The present analysis uses f-plane quasi-geostrophic dynamics, which among other restrictions typically requires that topographic amplitude, δh, scales as the assumed small Rossby number, i.e.

$$\frac{U}{f_0 L} = \varepsilon \sim \frac{\delta h}{H} \ll 1,$$

where H is a vertical length scale. This restriction will be avoided in the next section, but applies here. The use of an f-plane is easily justified; for example, Pickart (1995) argues the ratio of topographic to planetary β inshore of the Gulf Stream in a region of observed topographic waves grows to greater than 25.

Barotropic and two-layer baroclinic systems are considered; their governing equations are

$$\frac{\mathrm{d}}{\mathrm{d}t} q_i = \left(\frac{\partial}{\partial t} + \frac{J(p_i, \cdot)}{f_0} \right) q_i = 0, \tag{1}$$

where

$$q_i = \nabla^2 p_i + \frac{f_0}{H_1} h_b, \quad i = 1 \tag{2}$$

for the barotropic case, and

$$q_1 = \frac{\nabla^2 p_1}{f_0} - \frac{f_0}{g' H_1}(p_1 - p_2),$$

$$q_2 = \frac{\nabla^2 p_2}{f_0} + \frac{f_0}{g' H_2}(p_1 - p_2) + \frac{f_0}{H_2} h_b \tag{3}$$

for the baroclinic case. In the above, f_0 denotes the constant Coriolis parameter, H_i the thickness of layer i and g′ the reduced gravity of the interface. The quantities p_i and q_i represent layer i "reduced" pressure (i.e., pressure divided by a reference density ρ_0) and potential vorticity and J denotes the usual Jacobian operator. Topography, h_b, appears in this problem only in the lower layer equation (which in the barotropic fluid is the only layer) as a contribution to potential vorticity.

Neglecting lateral viscosity and bottom drag, (1) expresses the conservation of layer potential vorticity following fluid parcels, an exact solution of which is

$$q_i(x, y, t) = q_i(x_{i0}, y_{i0}, 0), \tag{4}$$

where

$$(x, y) = (x_{i0}, y_{i0}) + \int_0^t [u_i(x(t), y(t), t), v_i(x(t), y(t), t)]\, \mathrm{d}t. \tag{5}$$

Velocities $(u_i, v_i) = (-p_{iy}, p_{ix})/f_0$.

2.1. Barotropic analysis

Consider first the barotropic case. The basic analysis to follow is used several times in later sections and so is introduced here in some detail. Suppose this system is initially at rest and has a bottom topography defined by

$$h_b = \begin{cases} h_0, & y \geqslant a \\ h_0/2, & a \geqslant y \geqslant -a \\ 0, & -a \geqslant y, \end{cases} \tag{6}$$

i.e. the topographic form is meant to crudely represent a "slope", contained between $\pm a$, separating two abyssal plains, located at $|y| > a$. This general configuration will be referred to as "abrupt" and "confined", in that variations in the topography are of limited extent. The limit of $a = 0$ is that of a cliff face, or escarpment in topography with infinite slope. Of course, real ocean topography is seldom a cliff, but this may still be a useful idealization when topographic scales are short compared to dynamical fluid scales. For example, in Pickart (1995), the observed wave had an alongshore wavelength of roughly 300 km, while regional ocean depth changed by 3000 m in about 100 km. The effects of a finite slope, such as characterizes the Mid-Atlantic Bight, can also be examined in this framework by allowing for

separation between the cliffs (i.e. "a" is nonzero). A schematic of the system appears in Fig. 1.

If the otherwise resting fluid is perturbed, a nontrivial relative vorticity distribution will result (see Fig. 2). Assuming the amplitudes of the frontal disturbances are smaller than "$2a$", the pressure field will be governed by

$$\nabla^2 p_1 = \begin{cases} -f_0 q_0/2, & -a<y<L_2, \quad a<y<L_1 \\ f_0 q_0/2, & L_2<y<-a, \quad L_1<y<a \\ 0, & \text{elsewhere,} \end{cases} \tag{7}$$

where L_2 and L_1 denote the south and north fronts, respectively, and $q_0 = f_0 h_0/H_1$. The formal solution for pressure can be written

$$p_1(\vec{x},t) = -\int_{-\infty}^{\infty}\int_{a}^{a+L_1(\zeta,t)} \frac{f_0 q_0}{4\pi} \ln(|\vec{\zeta}-\vec{x}|)\,d\eta\,d\zeta - \int_{-\infty}^{\infty}\int_{a}^{a+L_2(\zeta,t)} \frac{f_0 q_0}{4\pi} \ln(|\vec{\zeta}-\vec{x}|)\,d\eta\,d\zeta \tag{8}$$

by Green's theorem. The variables $\vec{x} = (x,y)$ denote the observation points and $\vec{\zeta} = (\zeta,\eta)$ the source points. It is assumed in the above that $L_i(x,t)$ is a single valued function of x (although generalization to multi-valued functions of x is straightforward). This is certainly consistent with a linear wave analysis, but will also be used when the weakly nonlinear properties of the system are explored.

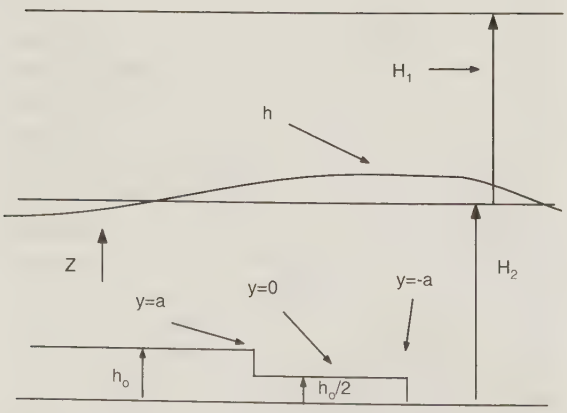


Fig. 1. Schematic of waves on weak topography. A two layer system is shown and bottom relief is modeled using escarpments separated by a distance $2a$. Both barotropic and baroclinic settings are considered; in the former, the subsurface interface is suppressed.

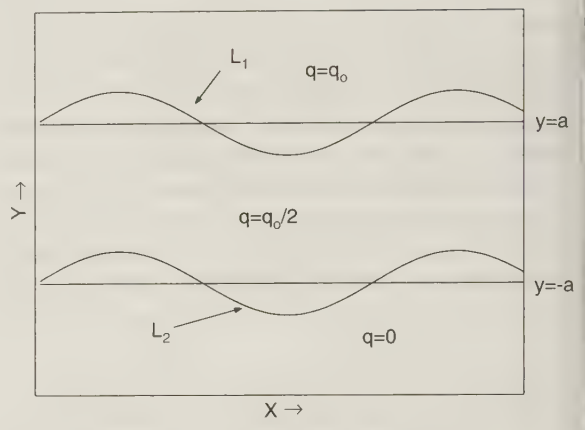


Fig. 2. Perturbing a resting fluid over the topography in Fig. 1 generates relative vorticity associated with the escarpments. The fronts, labelled "L_1" and "L_2", are shown in this schematic, along with the potential vorticity fields defining them.

The fronts are assumed to be material surfaces, i.e.

$$L_{it} + u(x, a_i + L_i, t)L_{ix} = v(x, a_i + L_i, t), \quad i = 1,2, \tag{9}$$

where $a_i = (-1)^{i+1}a$. Consider now the function $p_{L_i}(x,t) = p_1(x, a_i + L_i(x,t), t)$, i.e. pressure evaluated at a front. A little algebra shows (9) can in general be written

$$L_{it} = \frac{1}{f_0}(p_{L_i})_x, \tag{10}$$

an equation related to one derived in Pratt and Stern (1986) for long waves on a baroclinic jet.

In principle, (8) can be inserted in (10) to yield exact equations for the evolution of the fronts, although they require numerical solution. Part of the interest motivating the present study, however, is to develop an analytical appreciation of topographic variability. To serve this aim, (10) is approximated by

$$L_{it} = \frac{1}{f_0}\frac{\partial}{\partial x}(p_{L_i}(x,t)) \approx \frac{1}{f_0}\frac{\partial}{\partial x}\left(p_1(x,a_i,t) + \left[\frac{\partial}{\partial y}p_1(x,a_i,t)\right]L_i\right) + O((kL_0)^3), \tag{11}$$

where L_0 is a measure of the disturbance amplitude and k is its primary wavenumber. The above Taylor expansion for the frontal pressure is accurate as long as the wave steepness (kL_0) is small. The various derivatives of the pressure field at the fixed topographic break points can now be readily evaluated from (8), although they are as yet in integral form.

The wave characteristics of this system can be extracted by retaining only the leading order contribution to (11). If the resulting equation is then Fourier transformed in the observation variable x, and the dependence of the Green's function on the separation of the source and observation points is exploited, one obtains

$$\hat{L}_{jt} = \frac{q_0 i\,\mathrm{sign}(k)}{4}\int_{-\infty}^{\infty}\left(\int_{a_1}^{a_1+L_1} e^{-2\pi|k(a_j-\eta)|}\,d\eta + \int_{a_2}^{a_2+L_2} e^{-2\pi|k(a_j-\eta)|}\,d\eta\right)e^{-2\pi ik\zeta}\,d\zeta, \qquad (12)$$

where the "hat" ($\hat{}$) denotes a Fourier transformed variable. The result of integrating (12) in η involves $e^{-2\pi|kL_i|}$, which for small wave steepness can be expanded by Taylor series. Retaining the leading order term in L_i only eventually leads to

$$\hat{L}_{1t} = \frac{q_0 i\,\mathrm{sign}(k)}{4}\hat{L}_1 + \frac{q_0 i\,\mathrm{sign}(k)}{4}e^{-4\pi|k|a}\hat{L}_2,$$
$$\hat{L}_{2t} = \frac{q_0 i\,\mathrm{sign}(k)}{4}e^{-4\pi|k|a}\hat{L}_1 + \frac{q_0 i\,\mathrm{sign}(k)}{4}\hat{L}_2. \qquad (13)$$

Assuming $L_{it} \sim e^{2\pi i\omega t}$ yields a quadratic equation for the two eigenfrequencies

$$2\pi\omega_\pm = \frac{q_0}{4}\mathrm{sign}(k) \pm \frac{q_0}{4}\mathrm{sign}(k)e^{-4\pi|k|a} \approx \frac{q_0}{2}(\mathrm{sign}(k) - 2\pi ka, 2\pi ka), \qquad (14)$$

where it has been assumed in the latter formula that $|k|a \ll 1$.

These eigenvalues imply dispersion relations with some rather odd properties. Note that the first mode (ω_+) depends only on the escarpment (i.e. it exists even if $a = 0$), while the second mode (ω_-) requires a "slope" (i.e. $a \neq 0$)[1]. The escarpment case ($a = 0$) corresponds to the solution discovered by Rhines (1969) for barotropic double Kelvin waves. The surviving mode in this case is seen to have constant frequency, implying group velocities ($\partial\omega/\partial k$) vanish for all wavenumbers. Phase speeds, however, are nontrivial, and directed to the west.

The inclusion of a finite slope introduces dispersion to ω_+ (see (14)). At long wavelengths, the group velocity is constant. This would normally be associated with shape preserving, nondispersive behavior, but the picture here is complicated by the differing phase speed and group velocities. Note also that the group velocities of this mode are pseudo-eastward over the whole wavenumber regime. The eigenvector associated with this eigenvalue yields $L_1 = L_2$, i.e. the fronts are "in-phase" and thus sinuous.

The other ω_- mode, dependent upon finite slope ($a \neq 0$), is quite distinct in properties. It exhibits classical nondispersive behavior (i.e. $c_p = c_g$) in the long wave limit and is thus much more like typical planetary waves. Most importantly, these wave possess pseudo-westward directed group and phase speeds throughout their wavenumber regime. The eigenvector associated with their eigenvalue shows that $L_1 = -L_2$ and thus that the fronts are "out-of-phase", or "varicose".

Indeed, it is the distinct frontal structure of the waves that leads to their distinct dispersion relations. In-phase fronts reinforce, so the longest waves have the highest frequencies. Out-of-phase fronts interfere, and shorter waves thus increase in frequency. The two dispersion relations converge to

$$\omega_\pm \to \frac{q_0}{4\pi}\mathrm{sign}(k) \qquad (15)$$

as the short wave limit is approached ($ka \to \infty$) implying for weak slopes that each front behaves as an independent double Kelvin wave.

Nonlinear effects at weak amplitude can be investigated by including the next order correction in wave steepness in the Taylor expansion of $p_L(x,t)$. This analysis can be carried out for the case shown in Fig. 1, i.e., for $a \neq 0$ (the results are shown in the appendix), but the results are

[1]Here the term "slope" is used qualitatively; the dynamics introduced by a is the interaction between two Rhines waves.

strongly, and unrealistically, constrained by the practical requirement that the front amplitude be less than the inter-cliff separation 2a. (Without this constraint, the integral formulas for the pressure do not generally yield analytically tractable forms.) We therefore continue the weakly nonlinear part of this analysis, and indeed of all the following cases, examining the special case $a=0$. The advantage is that the disturbance amplitude is clearly not restricted as before, but we will consider only the nonlinear behavior of the sinuous mode. The only result of the nonlinear analysis for nonzero a that will be used later is the (obvious) statement that the varicose and sinuous modes are coupled.

Pressing on with $a=0$, there are two sources in (11) of nonlinearity: the obvious one is the quadratic term, i.e. $p_{1y}L$ (where the subscript "i" has been suppressed), but it is matched in amplitude by the second order contribution to p_{1x}. A concise way to illustrate the nonlinear behavior is to write the frontal evolution equation in physical space, which can be done by combining the nonlinear terms with the inverse Fourier transform of the dispersion relation. The result is

$$L_t - \frac{q_0}{8}(|L|L)_x + \frac{q_0}{2}\int_{-\infty}^{\infty} \ln(|u-x|)L_u \, du = 0. \tag{16}$$

The properties of this equation are interesting. The nonlinear term is a mild variant of the familiar amplitude dependent steepening of the KdV equation. The absolute value $|L|$ makes the effect symmetric in the cross topography direction, as is seen upon the substitution $L \to -L$. The interesting characteristic of the dispersive tendency is that it is nonlocal, i.e., net dispersion at a point is affected by frontal slopes at distant locations (although clearly local dispersion L_x has the largest effect).

The preceeding analytical results were tested by using a barotropic QG grid-point numerical model with a smooth, rather than discrete, bottom topography. Specifically, the total fluid depth was 1000 m and the net bottom variation was set to either 50 or 100 m and distributed as a hyperbolic tangent with various decay scales. While these are relatively mild bathymetric perturbations, they still generate strong potential vorticity. The 50 m case with a 5 km lateral scale has a maximum topographic β 50 times the planetary one.

A variety of initial potential vorticity fields were tested. Pure sinusoidal disturbances at given wavelengths to the resting potential vorticity field supported the dispersion relation of both modes. Other experiments used Gaussian initial disturbances to the potential vorticity. Results from such an experiment appear in Figs. 3 and 4. The net topographic step here was 50 m with a meridional decay scale of 20 km. The model employed a hyperviscosity coefficient of $-3\times10^{-8}\,m^4/s$ and horizontal resolutions of 5 km. Reentrant boundary conditions were applied on the east and west ends of the channel. The domain dimensions were 2560 km zonal by 1000 km meridional. The amplitude of the disturbance was 40 km and it had a zonal e-folding scale of 165 km.

Plan views of potential vorticity at various times are shown in Fig. 3 and a time longitude plot of the potential vorticity at the topographic break ($y=0$) appears in Fig. 4. The high frequency oscillations apparent in the latter plot are implicit in the former plot. Note that both plots show the timescales for the packet to migrate are long relative to the oscillation period. This is consistent with the weak group velocity of the dispersion relation relative to the phase speed (see (14)). The primary migration of the wave packet is to the east (indicated by a solid line in Fig. 4). The speed of the movement is about 0.04 m/s, and comparable to the predicted linear eastward group velocity for a slope of 20 km width. At day 120 in Fig. 3, evidence of a second, varicose, mode is seen. The track of this mode is indicated in Fig. 3 by the second solid line, beginning at roughly day 200. This pattern moves to the west at a speed of about 0.025 m/s, in agreement with the expected group speed. The late generation of this mode indicates it is the product of nonlinear excitation due to the nonlinear coupling between the wave modes.

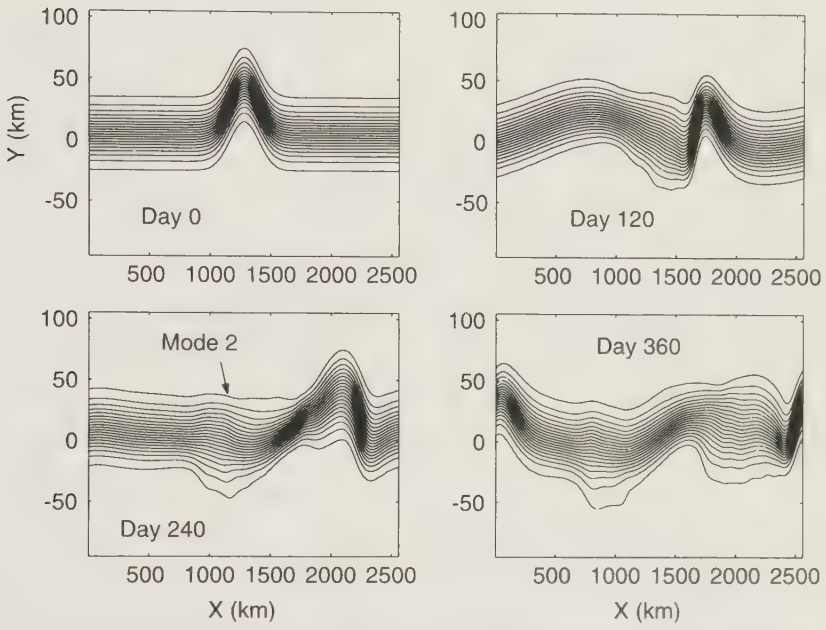


Fig. 3. Plan views of potential vorticity in a barotropic fluid with an escarpment. The otherwise resting fluid has been perturbed by a sinuous Gaussian shoreward displacement of 165 km width. The initial condition and results at days 120, 240 and 360 are shown. Note the primary eastward migration of the signal. The location of a varicose disturbance is labelled for day 240.

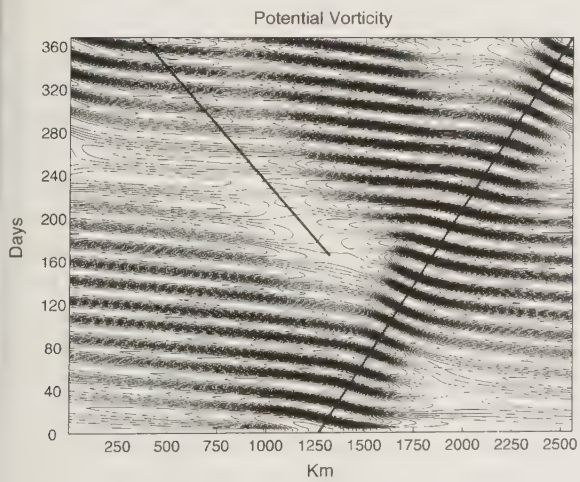


Fig. 4. Time-longitude plots of potential vorticity at the topographic break ($y=0$) from the experiment shown in Fig. 3. Solid lines trace the sinuous mode eastward propagation and the westward varicose mode propagation.

2.2. Baroclinic analysis

Consider now a two layer fluid over the topography (6). In this case, the solution in integral form can be written as

$$p_b(\vec{x},t) = \frac{-q_0 f_0 H_2}{4\pi}\int_{-\infty}^{\infty}\sum_{i=1}^{2}\left[\int_{a_i}^{a_i+L_i}\ln(|\vec{\zeta}-\vec{x}|)\,\mathrm{d}\eta\right]\mathrm{d}\zeta,$$

$$p_c(\vec{x},t) = -\frac{q_0 f_0}{4\pi}\int_{-\infty}^{\infty}\sum_{i=1}^{2}\left[\int_{a_i}^{a_i+L_i}K_0\left(\frac{|\vec{\zeta}-\vec{x}|}{R}\right)\mathrm{d}\eta\right]\mathrm{d}\zeta, \tag{17}$$

where K_0 is the usual modified Bessel function, $H=H_1+H_2$, $p_b=(H_1p_1+H_2p_2)$, $p_c=(p_1-p_2)$, $q_0=\frac{f_0h_0}{H_2}$ and other notation is as defined earlier. Pressures in each layer are given by $p_1=\frac{p_B+H_2p_c}{H}$ and $p_2=\frac{p_B-H_1p_c}{H}$ and the evolution of the fronts is governed by

$$\frac{\partial}{\partial t}L_i=\frac{1}{f_0}\frac{\partial}{\partial x}p_2(x,a_i+L_i,t).$$

The linear properties of the system can be found following the procedure outlined in the previous section and yield the dispersion

relationships

$$\omega_{\pm} = \frac{q_0 H_2 \mathrm{sign}(k)}{8\pi H}(1 \pm \mathrm{e}^{-4\pi|k|a}) + \frac{q_0 H_1 k}{4H\sqrt{R^{-2}+4\pi^2k^2}} \times \left(1 \pm \mathrm{e}^{-2a\sqrt{R^{-2}+4\pi^2k^2}}\right) \tag{18}$$

provided that wavelengths are long compared to wave amplitude. As in the barotropic case, two eigenfrequencies are found, with the ω_+ mode connected to "in-phase", sinuous oscillations of the front and the ω_- mode connected to "out-of-phase", varicose oscillations.

The former modal frequency converges to

$$\omega_+ \to \frac{q_0 H_2}{2\pi H}\left[\frac{\mathrm{sign}(k)}{2} + \frac{H_1 \pi k R}{H_2\sqrt{1+4\pi^2(kR)^2}}\right] \tag{19}$$

in the limit $a \to 0$, a dispersion relation found by Rhines (1977) and later by Willmott (1984). Clearly the first quantity in (19) comes from the barotropic properties of the system; the new dynamics caused by stratification are found in the second contribution. The form of the latter is the same as in Pratt and Stern (1986) for reduced gravity waves on a jet. Note that the escarpment waves in (19) are westward directed for all wavenumbers in both phase and group speeds

$$\frac{\partial}{\partial k}\omega = \frac{q_0 H_1}{2H}\frac{R}{(1+4\pi^2k^2R^2)^{3/2}}$$

but have the unusual property at long wavelengths that the constant, nondispersive group velocity differs from the phase speed.

Tainting the Rhines "double Kelvin Wave" with a finite bottom slope, however, results in quite distinct dispersion properties. For example, the sinuous mode group velocity with nonzero slope

$$\frac{\partial}{\partial k}\omega_+ = \frac{q_0 H_1 R}{4H(1+4\pi^2k^2R^2)^{3/2}}\left(1+\mathrm{e}^{-2a\sqrt{R^{-2}+4\pi^2k^2}}\right) - \frac{q_0 H_2 a}{2H}\mathrm{e}^{-4\pi|k|a} - \frac{q_0 H_1 a}{H} \times \frac{2\pi^2(kR)^2}{1+4\pi^2(kR)^2}\mathrm{e}^{-2a\sqrt{R^{-2}+4\pi^2k^2}} \tag{20}$$

(obtained from (18)) is quite articulated and typically exhibits a broad regime of eastward group velocity. This appears in Fig. 5, where several views of the modal dispersion relations appear. The upper panels illustrate the sinuous mode dispersion relation as a function of the separation parameter "a" and wavenumber. The deformation radius of the plot is 39 km and the layer thicknesses are both 1000 m. A global view appears on the left, from which it is apparent the character of the relation changes considerably near "a" = 10 km. Cross-sections for given "a" (labeled by their "a" value) appear on the right. The limiting $a = 0$ case shows a monotonic increase in ω with wavenumber, consistent with persistent westward drift. This is aggressively replaced by eastward group velocities at moderate to high wavenumbers for even small separations (e.g. $a = 5$ and 10 km). For separations comparable to the deformation radius, most wavelengths display eastward group velocity and at the greatest separations, long waves move rapidly east. Shorter waves, on the other hand exhibit westward phase speeds as the two escarpments begin to behave independently.

The tendency for eastward group velocity can be understood as the competition between the dynamics of the cliff and the slope. While the cliff face induces westward group velocity, the slope counters with an eastward tendency due to both barotropic and baroclinic dynamics. Depending on the ratio of slope to deformation radius, eastward effects can dominate, even at long wavelengths.

The varicose (ω_-) mode depends upon nonzero a for existence, as occurred for the barotropic case. An interesting distinction is the varicose mode propagates energy pseudo-westward throughout its wavenumber regime. This may be understood as the combined effect of the cliff and slope for the varicose mode, as both the vorticity and stretching dynamics result independently in westward motion. The varicose mode dispersion relation mode appears in the lower panels of Fig. 5; the presentation is analogous to that for the sinuous mode.

The nonlinear properties of the system can be examined by evaluating the next order in wave

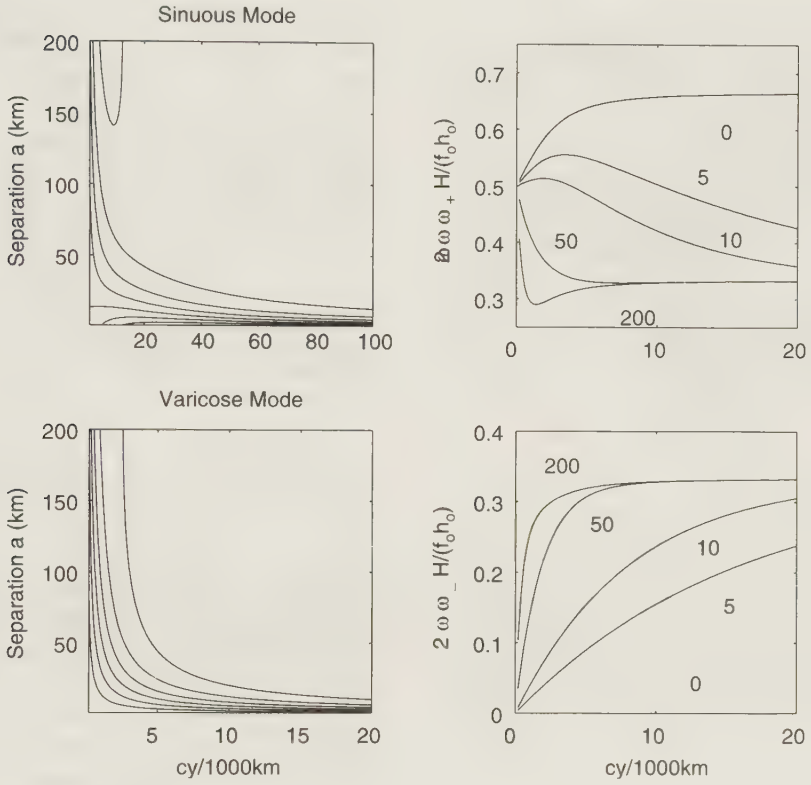


Fig. 5. The dispersion relation for the sinuous mode appears in the top panels and for the varicose mode in the bottom panels. The left-column contours frequency as a function of wavenumber and separation parameter ("a") and the right column shows frequency versus wavenumber for various values of "a". Note that eastward group velocity (frequency decreasing with wavenumber) typifies the sinuous mode while the varicose mode displays westward group velocity for all wavenumbers and separations.

steepness. Again, the focus is on the case of $a = 0$, which merges the two fronts and does not restrict frontal amplitude. One can easily write an integral form for the solutions analogous to (17) and evaluate the along-topography velocity at the break. There is a surprise here, in that the zonal flow is found in general to be

$$p_{2y}(x,0,t) = \frac{q_0 f_0}{2}|L| \tag{21}$$

and lacks explicit reference to the deformation radius. The second order contribution to $p_{2x}(x,0,t)$ also has this property. Aside from a factor of H_2/H, the steepening is identical to the barotropic case.

A wave equation for $a=0$ for this system accurate to $O(kL_0)^3$ can be written by assembling the various contributions discussed above, and is

$$L_t - \frac{q_0}{4}(|L|L)_x - \frac{q_0 H_2}{2\pi}\left[\frac{H_1}{H_2}\int_{-\infty}^{\infty} K_0\left(\frac{|x-p|}{R}\right)L_p\,\mathrm{d}p - \int_{-\infty}^{\infty}\ln(|p-x|)L_p\,\mathrm{d}p\right] = 0. \tag{22}$$

The shortwave limit yields an equation essentially like (16). For long waves, that equation is modified by a shape preserving nondispersive propagation.

Numerical solutions of a two-layer model in the presence of steep, but finite, bottom slopes have been examined in order to gauge the relevance of the preceeding analysis. Results from one particularly illuminating experiment are shown in Figs. 6 and 7. Here, the upper and lower layers have both been set to 1000 m, $g' = .03\,\mathrm{m/s^2}$ and the

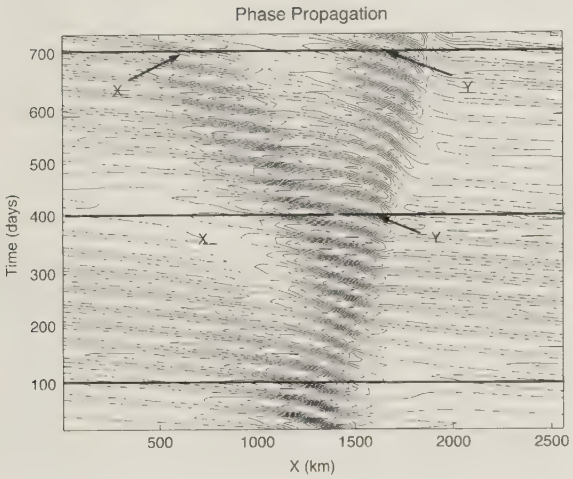


Fig. 6. Time-longitude plot of potential vorticity at the topographic break ($y=0$) from a two layer QG model with weak lower layer topography. The wave length scale and topographic lateral scale were chosen to mimic those of the Mid-Atlantic Bight setting considered by Pickart (1995). Note that the dominant propagation of the initial condition is eastward; later westward propagation is the product of nonlinear varicose mode excitation. The labeling and solid dark lines appearing at days 100, 400 and 700 correspond to the plan views in Fig. 7.

deformation radius $R = 39\,km$. The bottom consisted of a hyperbolic tangent profile with a net depth change of 50 m; the e-folding scale was chosen to be 30 km. The initial potential vorticity consisted of the resting state lower layer potential vorticity displaced northward in a 30 km Gaussian of a 120 km lateral e-folding scale. The result is an along escarpment disturbance of something like 300 km over a bottom shoaling over a length scale of roughly 100 km. Such wavelengths and lateral scales are like those described by Pickart (1995). The amplitude of the bottom change here is however considerably less than that occurring in the Mid-Atlantic Bight.

Fig. 6 shows a time-longitude plot of the potential vorticity field at the channel center. This figure resembles that shown earlier for the barotropic experiment, in that two oppositely propagating branches are seen. Plan views of the lower layer potential vorticity appear in Fig. 7. The times of these plots appear in the figure and are indicated in Fig. 6 by the bold lines. It is clear that the primary initial disturbance in this calculation is in the sinuous mode. The initial track in Fig. 6 shows eastward drift consistent with the previous analysis. Later, at about day 400, westward anomaly propagation appears. Referring to the plan view for that date, one sees evidence of a varicose mode. If the experiment is repeated with a lower amplitude for the initial disturbance (results not shown), westward propagation is largely absent. The interpretation of these results is that nonlinear interactions between the two modes excites the varicose mode.

3. Variability on abrupt, finite topography

The next problem to consider is that of finite amplitude topography, which is taken here to mean topography that the isopycnals intersect. This has traditionally been viewed as a regime that QG is unable to explore due to the parameteric restriction that $\delta h/H_2 \sim \varepsilon$. This restriction is avoided here by the technique of domain decomposition, i.e. the fluid is assumed to consist of two active layers south of the line $y=0$, and one active layer north of that line. Mysak et al. (1981) performed a linear stability analysis in such a system and Pedlosky (2000) has used a similar construct to compute planetary wave scattering off mid-ocean ridges.

The region north of the "shelf break" at $y=0$ has a total fluid depth of H_1-h_0 (where $h_0>0$). Assuming usual QG scaling, (1) for a single layer system is solved north of the break. South of the break, the equations for a two layer fluid are used. The isopycnal, nominally at a depth of $z=-H_1$, intersects the shelf break, and a nontrivial topographic potential vorticity field is presented to the upper layer due to the finite h_0. A schematic appears in Fig. 8.

We offer this, somewhat enhanced quasi-geostrophic system, as a dynamical model of variability in the presence of finite topography, although it is clearly a very idealized construct. A reviewer has voiced concern about whether the inherent simplifications of the model can be considered realistic. Ultimately, this framework will permit us to make

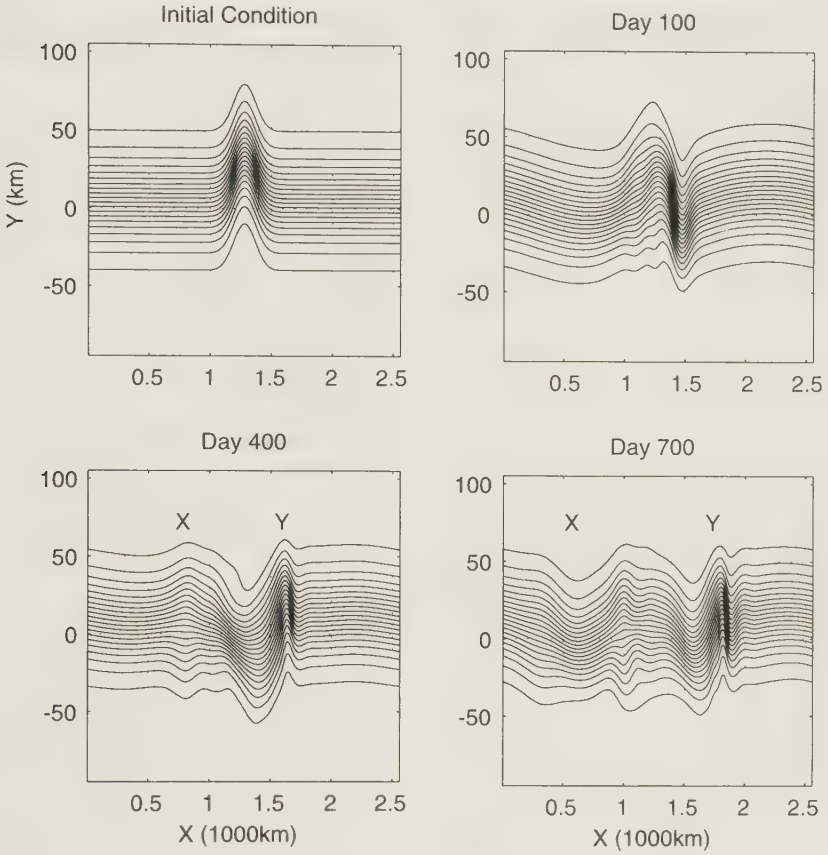


Fig. 7. Plan views of the potential vorticity field from the experiment in Fig. 6. The labels "X" and "Y" correspond to the labels there.

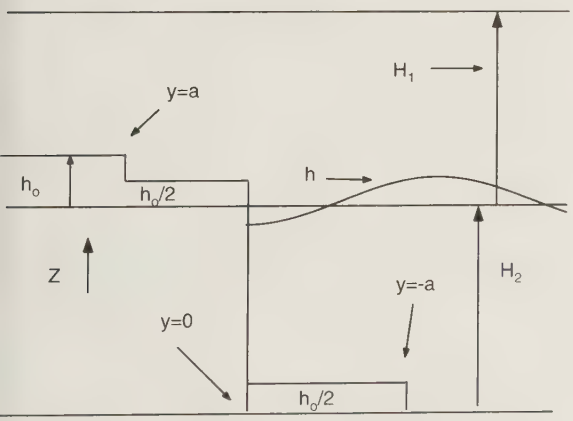


Fig. 8. Finite amplitude topography schematic. The isopycnal is assumed to strike the topography at a location $O(\varepsilon)$ below the cliff. The topography height anomaly is labelled $h_0/2$. At this location, the fluid changes from two layers to one layer. Also shown are two other escarpments at $y = \pm a$, one in the lower layer and one in the upper.

specific analytical statements about the expected behavior of topographically affected variability, and as little analytical progress in this area has been previously made, we suggest this supports the exploration of the model. Examination of the issues raised here by more advanced numerical models with topography should be pursued to support or refute our results.

At any instant, the upper and lower layers both have well-defined pressures and potential vorticities, so advancing the state in time is straightforward. The result of this is a prediction for potential vorticity, and the only subtlety involves inverting the potential vorticity into pressure. South of the break, the potential vorticity equations can be written in their modal forms for the barotropic and first baroclinic modes. North of the break, the only potential vorticity equation is of the barotropic form. In each of the subdomains,

computing a solution for pressure subject to given potential vorticity and proper boundary conditions is straightforward.

Consider now the boundary conditions that apply at the break. In the lower layer, there can be no normal flow through the shelf, so one boundary condition is

$$p_{2x} = \left(\frac{p_b - H_1 p_c}{H}\right)_x = 0. \tag{23}$$

The upper layer potential vorticity can be written globally as

$$\nabla^2 p_1 - \frac{f_0^2}{g' H_1} V(-y)(p_1 - p_2) = f_0 q_1 - \frac{f_0^2 h_0}{H_1} V(y), \tag{24}$$

where the usual Heaviside step function is represented by $V(y)$

$$V(y) = \begin{cases} 1, & y > 0 \\ 0, & y < 0. \end{cases} \tag{25}$$

If (24) is now integrated from $y = -\iota$ to $y = \iota$, where ι is "small", the jump constraint

$$p_{1y}(x, \iota, t) = p_{1y}(x, -\iota, t) + \mathrm{O}(\iota) \tag{26}$$

is obtained. Multiplying (24) by y, integrating the result over the same interval and using (26) then yields

$$p_1(x, \iota, t) = p_1(x, -\iota, t). \tag{27}$$

The above two constraints, implying continuity of pressure and tangential velocity at the break, when combined with the no-normal lower layer flow condition, provide enough boundary conditions to determine unique solutions for the pressure fields given potential vorticities. In this paper, the upper layer condition

$$p_{1x}(x, \iota, t) = p_{1x}(x, -\iota, t). \tag{28}$$

will always be employed in place of (27). The two are equivalent if q_1 is everywhere finite. The method by which the potential vorticity field is inverted to yield pressure, given these constraints, is discussed in Appendix B.

For $a = 0$, the front, L, is now obviously confined entirely to the upper layer and is governed by

$$L_t = \frac{1}{f_0}\frac{\partial}{\partial x} p_1(x, L, t) \approx \frac{1}{f_0}\frac{\partial}{\partial x}[p_1(x, 0, t) + p_{1y}(x, 0, t)L]. \tag{29}$$

The linear properties of the system can be extracted in this case in a manner analogous to that used in the previous sections. Fourier transforming (29) and performing some straightforward algebra yields the surprisingly simple wave dispersion relation

$$\omega = \frac{q_0 H k \,\mathrm{sign}(k)}{(H + H_1)2\pi k + \mathrm{sign}(k) H_2 \sqrt{R^{-2} + 4\pi^2 k^2}}, \tag{30}$$

which exhibits pure westward group velocity. There are imprints of both (14) and (19) in this dispersion relation, but the interesting result is that it is not a simple linear combination of the barotropic and baroclinic dispersion relations, as it was for weak topography (see (19)). Long waves ($kR \to 0$) converge to $\omega \to 2q_0 \pi k R H / H_2$, which displays classic nondispersive properties, e.g., $c_p = c_g$. In this property, it differs from either of the weak topography results, but in c_g is like the baroclinic system at long wavelengths. Thus finite topography shuts off the high frequency oscillation of long waves, ultimately a property deriving from barotropic dynamics. It does not interfere with the shape preserving dispersion tendencies that the weak topography system obtains from the baroclinic structure. These results can be understood as the result of the suppression of the barotropic mode by the finite amplitude topography. To see this, note that the lower layer no-flux condition implies

$$p_{bx} = H_1 p_{cx} \tag{31}$$

thus slaving the barotropic mode to the baroclinic mode. This combined with the expression for upper layer pressure at the shelf break shows

$$p_{1x} = p_{cx} \tag{32}$$

illustrating the filtering of barotropic dynamics.

In the short wave limit ($kR \to \infty$), the dispersion relation (30) converges to (14) where the active fluid layer is the shallow layer H_1. This result may be understood in the sense that short waves experience the deformation radius as a very large length scale, and hence the interface behaves essentially like a rigid surface. Such waves are also surface trapped as analysis of p_{2x} shows it vanishes like $(kR)^{-1}$. This result, while different in detail from the weak topography case (where the wave was bottom trapped) is nonetheless qualitatively similar; the wave is confined to the layer where the potential vorticity step exists.

The nonlinear wave properties of this system can be probed by following the procedure employed earlier. The Fourier transform of $p_{1y}(x,0,t)$ can be computed as described in Appendix B and is

$$\widehat{p_{1y}}(k,0,t) = \frac{f_0 q_0 \hat{L}_+}{2} \times \frac{(2H_1\pi k + H_2 \mathrm{sign}(k)\sqrt{R^{-2}+4\pi^2k^2})}{\left[(H_1+H)\pi k + H_2\frac{\mathrm{sign}(k)}{2}\sqrt{R^{-2}+4\pi^2k^2}\right]} - \frac{2H\pi k f_0 q_0 \hat{L}_-}{2(H+H_1)\pi k + \mathrm{sign}(k)H_2\sqrt{R^{-2}+4\pi^2k^2}}, \quad (33)$$

where $L_+(L_-)$ denotes the onshore (offshore) portion of the front. The above has purposely been truncated to $O((kL_0)^2)$, as only the lowest order of p_{1y} is required. The physical space representation of nonlinear effects requires an inverse Fourier transform of the above and of the quadratic contributions to p_{1x}. While possible, the result is a rather cumbersome set of convolution integrals. On the other hand, the limiting case of long waves ($kR \to 0$) is illuminating and demonstrates that nonlinearity in general distinguishes strongly between on and offshore disturbances. For long waves, the leading order approximation of the nonlinear terms in physical space reduces to the contribution

$$(L_+L_+)_x + (L^2)_x.$$

Note, steepening favors onshore perturbations and the onshore–offshore symmetry is broken. Interestingly, in the shortwave limit ($kR \to \infty$), nonlinear effects converge to the same dependence as the purely barotropic problem, i.e. $(|L|L)_x$. This tendency is onshore–offshore symmetric, and fits with the linear behavior of the short-wave limit discussed above. Onshore–offshore symmetry is exceptional however, as argued in Fig. 9. This plot shows the front tendency, L_t, in physical space compared to front structure for a 30 km wavelength. The quantity L_t (solid line) was computed by evaluating (29) in the wavenumber domain and Fourier transforming to physical space. The dashed curve is the front structure. Note that the front tendency depends strongly on the sign of the frontal displacement, and is characterized by shorter length scales on the offshore side.

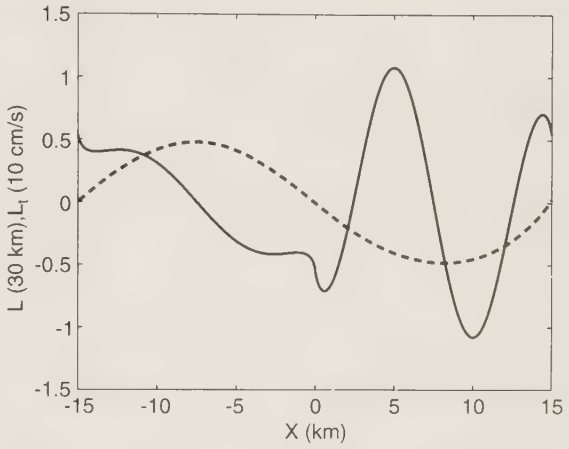


Fig. 9. Front tendency as a function of onshore–offshore wave amplitude. The dashed curve indicates the front position along the topographic break, the solid curve is the front tendency L_t computed by (29). Note that the tendency depends strongly on the sign of the front disturbance relative to the topographic break.

Given the presence of the steepening in this problem, an obvious question is if it can be balanced by linear dispersion. At long wavelengths, leading order dispersion is shape preserving, so the next order (p'_{1x}) must be obtained. The result is nonlocal, i.e.

$$p'_{1x} = \frac{-(H+H_1)Hf_0q_0R^2}{H_2^2\pi} \times \int_{-\infty}^{\infty} \ln(|x-u|)L_{uuu}(u)\mathrm{d}u, \quad (34)$$

which is a curious mixture of baroclinic (represented by R) and barotropic (represented by the ln function) dynamics. This effect is somewhat like the dispersion in the KdV equation, although the nonlocality is novel.

In summary, the equation governing long-wave disturbances in the present finite-amplitude topography system can be written as

$$L_t - \frac{q_0}{4}(L_+L_+)_x - \frac{q_0}{4}(L^2)_x - q_0 R\frac{H}{H_2}L_x = \frac{-(H+H_1)Hq_0R^2}{H_2^2} \times \frac{1}{\pi}\int_{-\infty}^{\infty} \ln(|x-u|)L_{uuu}(u)\mathrm{d}u. \tag{35}$$

The behavior expected from this equation is that onshore long waves should steepen and break toward the west, while offshore waves will steepen, but somewhat more weakly and break towards the east. As the length scale of the front decreases, dispersive effects will grow in amplitude.

Both barotropic and baroclinic weak topography settings demonstrated a sensitivity to slopes, in that group velocities in the presence of a slope were dominantly eastward for the sinuous mode. To check if the same sensitivity exists for finite amplitude topography requires that both on and offshore slopes be examined. These appear in Fig. 8 as the escarpments at $\pm a$. Including both escarpments and the primary escarpment at the shelf break yields a problem too cumbersome to be analytically useful; thus on and offshore slopes have been investigated separately (i.e. only one of the escarpments at $y = \pm a$ is retained). Of these two settings, the one that at least superficially resembles Pickart (1995) is the offshore slope, as that wave propagated toward the shelf from the deep water prior to paralleling the break. This setting is considered first, and as done previously, the focus is on the linear, dispersive properties of the system.

The same procedure can be applied to the finite topography/offshore slope problem as described earlier in this section, although the algebra involved in inverting the potential vorticities is considerably more involved. The result of the analysis is

$$\omega_\pm = \frac{\frac{q_0\,\mathrm{sign}(k)H}{D} + A}{4\pi} \pm \frac{\left(\left(\frac{q_0\,\mathrm{sign}(k)H}{D} - A\right)^2 + 4H_1H_2q_0^2M^2\right)^{1/2}}{4\pi}, \tag{36}$$

where

$$D = H + H_1 + \frac{H_2\mathrm{sign}(k)\sqrt{R^{-2}+4\pi^2k^2}}{2k\pi},$$

$$A = \frac{H_2\,q_0\,\mathrm{sign}(k)}{2H}(1-\mathrm{e}^{-4\pi|k|a}) + \frac{H_1\,q_0\,k\pi}{H\sqrt{R^{-2}+4\pi^2k^2}}\left(1-\mathrm{e}^{-2a\sqrt{R^{-2}+4\pi^2k^2}}\right) + \frac{q_0\,H_1H_2\,\mathrm{sign}(k)D}{H}M^2$$

and

$$M = \frac{\left(\mathrm{e}^{-2\pi|k|a} - \mathrm{e}^{-a\sqrt{R^{-2}+4\pi^2k^2}}\right)}{D}$$

It has been assumed that the topographic step amplitudes in the two layers are identical. These formulas, evaluated in the limit $a \to 0$, show the "+" mode corresponds to the single escarpment results in (30) while the "−" mode depends upon finite "a" for existence. The eigenvector structure associated with these eigenvalues can also be computed, but is no longer a simple constant relationship across all wavenumbers, as it was for weak topography. Last, the various limits (i.e. long wavelength, short wavelength and large separation a), all converge to sensible results. Characteristics of the dispersion relations appear in Fig. 10, from which it is seen that both modes now possess uniformly westward propagation for all separations a, and that the "+ (−)" mode corresponds to in (out) of phase frontal oscillations. This plot emphasizes that finite topography qualitatively affects variability, rather than quantitatively. The tendency for westward group velocities reflects the topographic suppression of the barotropic

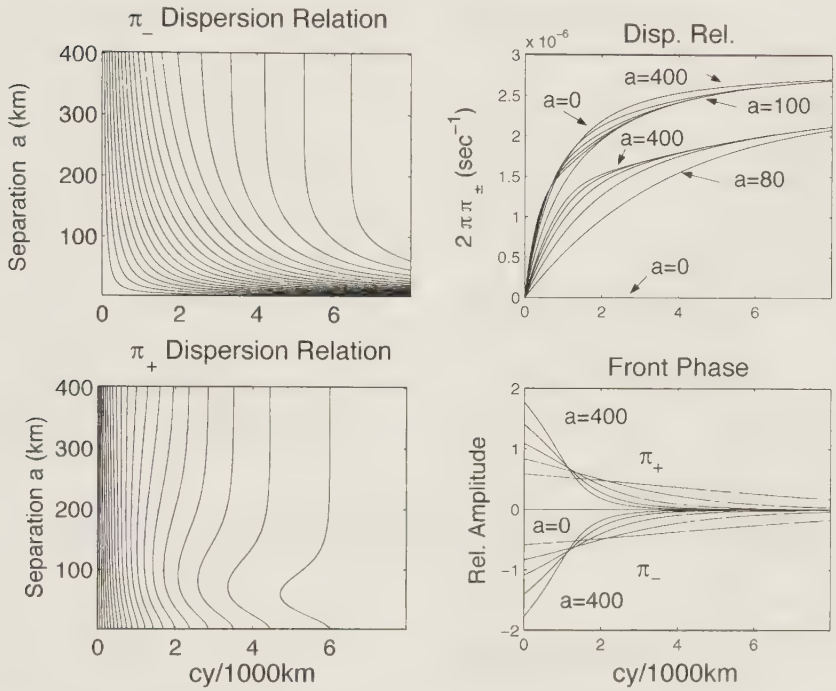


Fig. 10. Wave characteristics for an offshore slope with finite topography. The upper (lower) left panel displays the $\omega_{+(-)}$ dispersion relation as a function of wavenumber and escarpment separation. The upper right panel shows frequency against wavenumber for several separations and for both dispersion relations. The ω_- values are generally smaller. The lower right panel shows the wave structure of the eigenmodes. The ω_+ waves are all in phase and lie in the upper part of the diagram. There, the ratio of the front at the break to the front offshore is plotted. The ω_- waves are all out of phase and appear in the lower half of the diagram. There, the ratio of the offshore front to the front at the topography is plotted. Results for several values of the separation parameter are given. Note, for large separations, the fronts become decoupled, except at the longest wavelengths.

dynamics alluded to before. In this case, a degree of barotropic response is driven by the offshore escarpment, but the requirement of no-normal flow on the topography sets its strength so that baroclinic dynamics control the wave.

Onshore fronts differ slightly in character. The dispersion relations are found by solving the quadratic equation

$$\left(\frac{\omega}{q_0}\right)^2 - \left(\frac{\pi k}{\sqrt{1+4\pi k^2R^2}+2\pi kR} + A\right)\frac{\omega}{q_0} + \frac{\pi kA}{\sqrt{1+4\pi^2k^2R^2}+2\pi k} - \frac{e^{-4\pi|k|a}\pi^2k^2R^2}{\left(\sqrt{1+4\pi k^2R^2}+2\pi kR\right)^2} = 0,$$

where

$$A = \frac{1}{4} - e^{-4\pi|k|a}\frac{\left(\frac{1}{4} - \frac{\pi kR}{2\sqrt{1+4\pi^2k^2R^2}}\right)}{1+\frac{2\pi kR}{\sqrt{1+4\pi^2(kR)^2}}}$$

and a reduced gravity assumption has been made. Evaluating the formulas for $a\to 0$ shows the "+" mode corresponds to (30). The character of these waves appears in Fig. 11. The left-hand column contours frequency as a function of wavenumber and separation, while the upper right plot displays sinuous frequency versus wavenumber for several separations. The lower right plot describes modal structure; sinuous (varicose) modes are in (out of)-phase and thus in the upper (lower) diagram half. The amplitude ratio of the

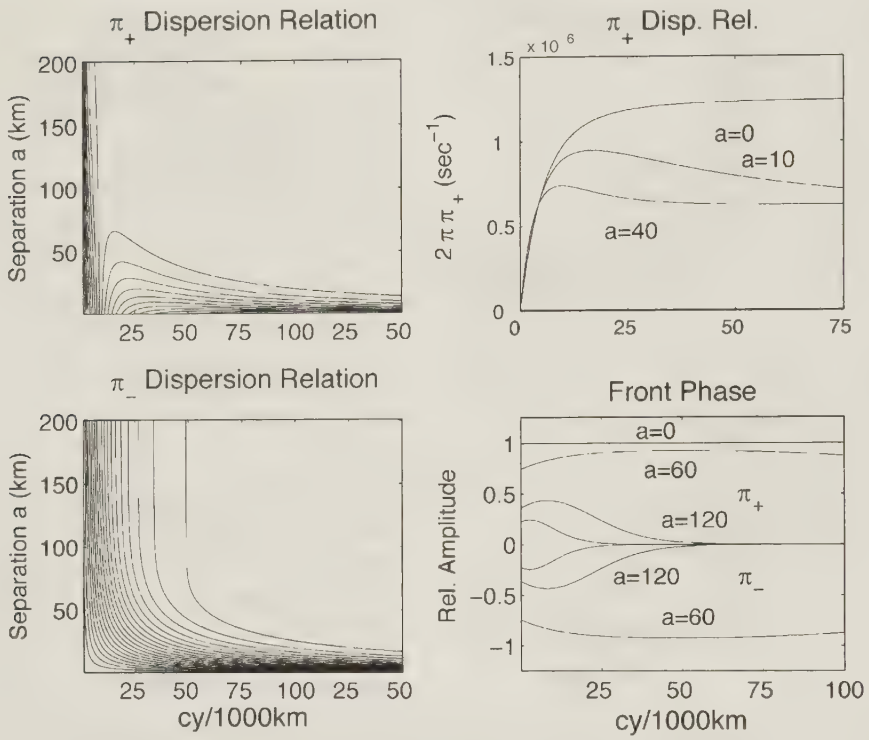


Fig. 11. Dispersion characteristics for an inshore escarpment in the presence of finite topography. The left column contours modal frequency as a function of wavenumber and separation. The upper right figure shows ω_+ as a function of wavenumber for several separation parameters. The $\omega_{+(-)}$ mode is sinuous (varicose) as shown in the lower right plot. The sinuous mode shows a weak tendency for eastward group velocities.

fronts is a strong function of wavenumber and separation.

Note the sinuous modes are not characterized by a uniformly westward group velocity in this case. On the other hand, eastward group velocities are weak and generally relegated to high wavenumbers where interaction between the escarpments is weak. Longer wavelengths, where the fronts do interact tend strongly to westward group velocities, sometimes with extreme values.

The behavior of this system was numerically examined by integrating the weakly nonlinear wave equation (29). Namely, at each time step the frontal path L was broken into L_+ and L_-, Fourier transformed in space and (45) was evaluated. An inverse Fourier transform then yielded both $p_{1x}(x,0,t)$ and $p_{1y}(x,0,t)$, from which an estimate of the front location at the next time could be obtained. Note this technique is more general than the frontal path equation (35) which is restricted to disturbances long compared to the deformation radius, thus enabling a variety of initial conditions to be tested.

We show in Fig. 12 the results of such an integration for a 350 km initial disturbance in a two layer fluid with deformation radius 20 km. The mean layer depths were 1000 m and the wave amplitude was 15 km. This wavelength is like that observed by Pickart (1995). The early evolution is primarily for the inshore meander to steepen to the west, as seen in days 30 and 60. This is consistent with the noted tendency for nonlinearity to be stronger for inshore meanders. For this wave amplitude, the increase in linear dispersion associated with the decrease in onshore length scale is apparently adequate to prevent wave breaking. The largest inshore meander is a persistent feature

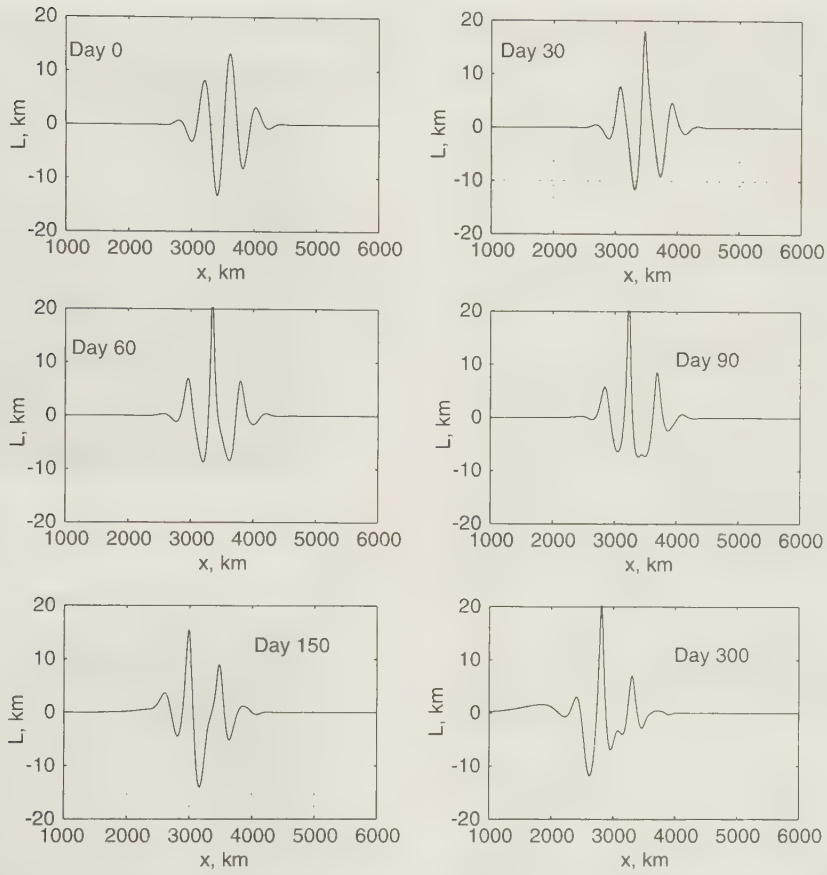


Fig. 12. An example of front propagation subject to finite amplitude topography. The weakly nonlinear front equation (29) was solved for an initial disturbance with primary length scale 350 km and an envelope of 1000 km. The deformation radius for this calculation was 20 km. Results from days 0, 30 60, 90, 150 and 300 are shown and emphasize the onshore/offshore asymmetry of the evolution coupled with westward propagation.

f the wave evolution, suggesting the system is apable of solitary wave-like behavior. On the ther hand, we have found at larger amplitudes nat the wave equation breaks down, signaling the nset of wave breaking. The behavior of the ffshore meander is different, with the develop- nent of a broad offshore trough at day 90. The verall propagation tendency for the packet is to ne west at roughly 0.02 m/s. Interestingly, even vith the relatively small deformation radius elative to wavelength, the waveform exhibits ehavior fairly far removed from the longwave mit. This is due to the slow dependence of the ispersion relation on wavenumber.

4. Discussion

Variability on steep and confined topography has been considered within the confines of an analytical, weakly nonlinear theory. Restricting the topography to finite numbers of escarpments permits the exact solution to be written in integral form, which can then be approximated to yield dispersion relations and evolution equations. The method can be used to examine a variety of topographic shapes, from finite slopes to finite topography characterized by isopycnal-bathymetric intersection.

To summarize the barotropic and baroclinic results for weak topography, some interesting modifications to classical GFD wave behavior result. The waves divide into those dependent upon the net depth change and those dependent upon the slope. The first type exhibits rapid oscillations due to barotropic effects, typically possess eastward group velocities dependent upon the slope and are sinuous in structure. The second, varicose modes require the slope to exist, and display the more classical tendency for westward group velocity. Interestingly, this is driven by both the barotropic and baroclinic dynamics of the waves. For finite topography, the phase and group velocities are typically westward for both sinuous and varicose modes, an effect due to the tendency for finite topography to filter barotropic dynamics. Also, such meanders evolve differently depending on whether they are on or offshore.

To place these results in an oceanic context, the majority of reported topographic wave observations are consistent with pseudo-westward propagation of energy. This appears in Pickart (1995) and in all of the regional studies that preceded it. We have also examined satellite observed SST from the Mid-Atlantic Bight in an attempt to characterize cross-shelf variability and propagation. An example of the results so obtained appears in Fig. 13, which compares two SST composites from Fall, 1998 in a region offshore of the northeast U.S. seaboard. The heavy black line indicates an SST front, and the arrows inshore and offshore frontal displacements displaying a wave-like appearance and southwestward, along-topography propagation at approximately .02 m/s. The location of these meanders is near the location of the site D mooring, whose topographic variability was analyzed originally by Thompson and Luyten (1976). The location of the array analyzed by Pickart is nearer the Gulf Stream exit point, nominally at (74W, 35.5N).

The cross-topography structure of Pickart's waves were described in Pickart and Watts (1990). There, they argue from current meter data near the 3000 and 4000 m isobaths that inshore and offshore wave edges are out-of-phase. Specifically, they find a 16-day lag between their current meter moorings C3 and C5 (situated just offshore

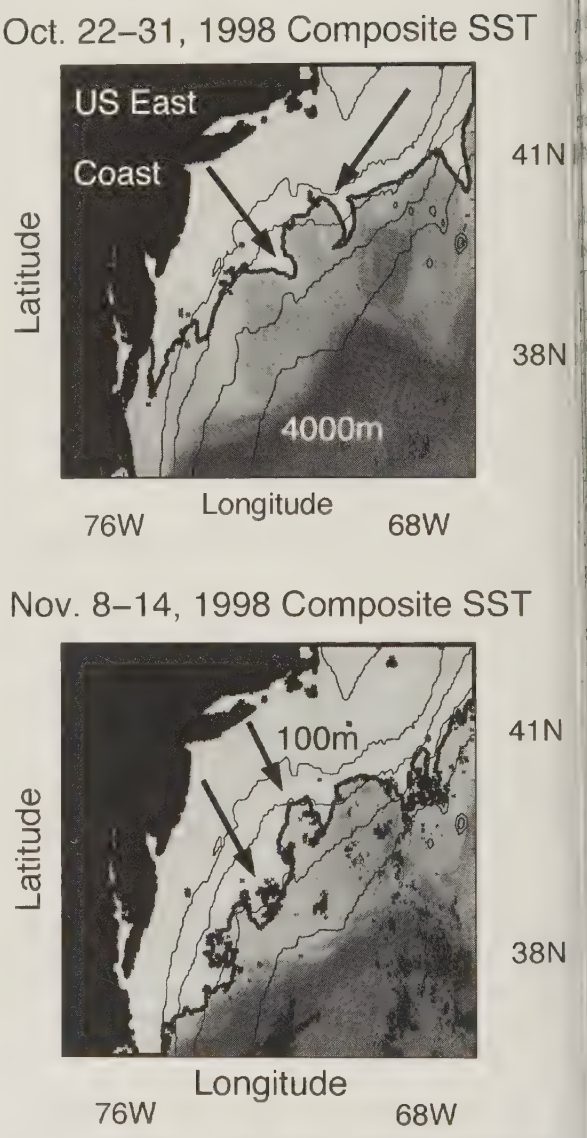


Fig. 13. SST just off the US east coast from late October 1998 is shown. The locations of the 4000, 3000, 2000 and 100 m topographic contours are shown, along with the outline of the coast. The heavy dark line accents an SST front and the arrows point at SST anomalies that have a sinuous wave-like appearance.

of 3000 m and onshore of 4000 m depths, respectively) in a region with dominantly 40 day variability. They interpret this as evidence of downslope phase propagation, which is in keeping with expectations built on topographic waves in a sloping bottom, QG, unbounded model. We point

ut here that the lag may also be interpreted as upporting a varicose wave structure, which uggests that the present analysis is an oceanically elevant technique for including topography and oundaries into a wave theory.

The meanders appearing in Fig. 13, on the other and, appear to be more properly described as inuous and yet display a pseudo-westward drift. his is a combination of behaviors that the present nalysis would ascribe to the effects of finite opography, again recommending the validity of he present approach. We do admit however that SST data has proven difficult to extract much nformation from. Attempts on our part to dentify finite amplitude effects using regional SST data have not been conclusive. This is due o the complicating observational factors of poradic cloud cover and seasonal SST migration, nd the competing effects of the main Gulf Stream urrent and the warm core rings that it sheds (two f which appear in the figure). Examining other lata sources for variability characteristics in egions of strong topography would be quite nteresting.

The structural change caused by finite topography leading to distinct behavior is its tendency to suppress barotropic dynamics. This reflects in the present analysis the no-flow boundary condition on deep layers that the finite topography enforces. Other curious ramifications of the topography include rather major modifications of nonlinear behavior. Finite amplitude effects discriminate between on and offshore sides with regards to strength, sign and lengthscale. The clearest example of this is seen in the longwave limit, where symmetric steepening for weak topography is replaced by opposite signed steepening depending on location (onshore forward breaking, offshore backward breaking) and enhanced onshore strength. At smaller scales, the onshore offshore differences are even more profound, possessing distinctions in the scales they impose on the wave and the sense of the steepening.

This analysis has considered very simple models of topographic variability in isolation of many potentially significant effects, like background currents, vortices and complex, along-shore topography. All of these features are evident in Fig. 13. Given that the nature of the variability in this theoretically simplified setting is modified at a fundamental level by finite topography, consideration of these other real world motivated effects becomes of interest. Such studies are the focus of on-going work.

Acknowledgements

WKD is supported by NSF grants OCE-0220700 and OCE-0220884. Peter Cornillon is gratefully thanked for his assistance in our examination of the Pathfinder SST data, and J.C. McWilliams is recognized for an interesting conversation at the early stages of this work. Jane Jimeian's assistance with various computational issues is greatly appreciated.

Appendix A. Nonlinearity for nonzero escarpment separation

The quantity $p_{1y}(x, a_i, t)$ can be computed directly from (8)

$$p_{1y}(x,0,t) = -\frac{q_0 f_0}{4\pi}\int_{-\infty}^{\infty}\int_{a}^{a+L_1}\frac{\eta}{|\vec{\zeta}-\vec{x}|}\,\mathrm{d}\eta\,\mathrm{d}\zeta - \frac{q_0 f_0}{4\pi}\int_{-\infty}^{\infty}\int_{-a}^{-a+L_2}\frac{\eta}{|\vec{\zeta}-\vec{x}|}\,\mathrm{d}\eta\,\mathrm{d}\zeta.$$

Fourier transforming the above equation in x, employing the formula

$$\int_{-\infty}^{\infty}\frac{1}{|\vec{\zeta}-\vec{x}|}\,\mathrm{e}^{-2\pi i k(x-\zeta)}\mathrm{d}(x-\zeta) = \frac{\pi}{|\eta-y|}\,\mathrm{e}^{-2\pi|k(\eta-y)|}$$

and expanding assuming small amplitude disturbances ($L_i \to 0$) eventually yields

$$\widehat{p_{1y}}(x,a,t) = \frac{f_0 q_0}{4}\widehat{|L_1|} - \frac{q_0}{4}\mathrm{e}^{-4\pi|k|a}\hat{L}_2 \tag{37}$$

along with an analogous form for $\widehat{p_{1y}}(x,-a,t)$. The next order corrections to $\widehat{p_{1x}}$ are easily seen to be

$$\widehat{p_{1x}}(x,a,t) = -\frac{q_0 f_0 \mathrm{sign}(k) i}{8}(L_1\widehat{|L_1|})_x + \frac{q_0}{8}\mathrm{e}^{-4\pi|k|a}(\widehat{L_2^2})_x. \tag{38}$$

The most important aspect of these formulas is that they nonlinearly link the fronts.

Appendix B. Potential vorticity inversion in the presence of finite topography

In the present problem, it is particularly convenient to employ a variant of the so-called capacitance matrix method (Cummins and Mysak, 1988) to enforce the conditions at $y=0$ on potential vorticity inversion. The scheme is based on the formal integral solutions to the potential vorticity equations like those used in the main body of the paper, where the integrals are performed over the regions where potential vorticity values are specified. In the capacitance matrix method, the domain of interest is artificially enlarged, and the potential vorticity field is augmented to include an artificial contribution that enforces the boundary conditions.

To illustrate the technique, consider the shallow fluid north of the break, defined over the half plane $y>0$. The domain of the shallow fluid is artificially enlarged to include the southern half plane, and the potential vorticity field is written

$$Q_1 = q_1 V(y) + \theta_1(x,t)\delta(y-y_1), \tag{39}$$

where δ denotes the Dirac delta function and $V(y)$ the Heaviside step function. Thus the enlarged q field, Q_1, consists of the true potential vorticity north of the break (q_1) and an ad hoc contribution at latitude y_1, where y_1 is necessarily chosen to lie in the artificial southern half plane. Similarly, the two layer fluid, lying literally in the southern half plane, is also artificially enlarged to include the northern half plane, where potential vorticity sources θ_c and θ_b are assumed to exist. Similar formulae are written for the barotropic and baroclinic potential vorticities, i.e.

$$\begin{aligned} Q_c &= q_c V(-y) + \theta_c(x,t)\delta(y-y_2), \\ Q_b &= q_b V(-y) + \theta_b(x,t)\delta(y-y_2), \end{aligned} \tag{40}$$

where the latitude $y_2>0$. The general solution for each mode is now written as a Green's function integral and derivatives denoting the velocities of the flow are evaluated at the break. Enforcing the three conditions at the break then yields three integral constraints governing the unknown potential vorticity enhancements θ_1, θ_c and θ_b.

Consider a single escarpment of amplitude h_0 (potential vorticity anomaly $q_0 = \frac{f_0 h_0}{H_1}$) at the topographic break. The condition (23) becomes

$$\begin{aligned} &\int_{-\infty}^{\infty} \theta_b(\zeta, y_2)\frac{\partial}{\partial x} G_b(\vec{x};\vec{\zeta})\mathrm{d}\zeta \\ &\quad - H_1 \int_{-\infty}^{\infty} \theta_c(\zeta, y_2)\frac{\partial}{\partial x} G_c(\vec{x};\vec{\zeta})\mathrm{d}\zeta \\ &\quad = H_1 f_0 q_0 \int_{-\infty}^{\infty}\int_{L_-}^{0} \frac{\partial}{\partial x} G_b(\vec{x};\vec{\zeta})\mathrm{d}\vec{\zeta} \\ &\quad + H_1 f_0 q_0 \int_{-\infty}^{\infty}\int_{L_-}^{0} \frac{\partial}{\partial x} G_c(\vec{x};\vec{\zeta})\mathrm{d}\vec{\zeta}, \end{aligned} \tag{41}$$

where $G_{b,c}$ represents barotropic and baroclinic Green's functions and L_- denotes those parts of the front south of the break. Note that only L_- appears in (41); both northern (L_+) and southern front components appear in the other two constraints although each occurs independently of the other. The front, L, is now obviously confined entirely to the upper layer and is governed by

$$\begin{aligned} L_t &= \frac{1}{f_0}\frac{\partial}{\partial x} p_1(x, L, t) \\ &\approx \frac{1}{f_0}\frac{\partial}{\partial x}[p_1(x,0,t) + p_{1y}(x,0,t)L]. \end{aligned} \tag{42}$$

Comparable equations to (41) can be written for (26) and (28). Now consider a Fourier transform of (41) in the observation variable x. The Green's functions depend only on the separation between source and observation points, so (41) becomes

$$\begin{aligned} &\frac{\widehat{\Theta_b}\,\mathrm{sign}(k)}{2} - \frac{kH_1\pi}{\sqrt{R^{-2}+4\pi^2k^2}}\widehat{\Theta_c} \\ &\quad = \frac{H_1 f_0 q_0}{2}\int_{-\infty}^{\infty}\int_{L_-}^{0} \mathrm{e}^{2\pi|k|\eta}\,\mathrm{sign}(k)\,\mathrm{d}\eta\mathrm{e}^{-2\pi ik\zeta}\,\mathrm{d}\zeta \\ &\qquad + \frac{H_1 q_0 \pi k f_0}{\sqrt{R^{-2}+4\pi^2k^2}} \\ &\qquad \times \int_{-\infty}^{\infty}\int_{L_-}^{0} \mathrm{e}^{\eta\sqrt{r^{-2}+4\pi k^2}}\,\mathrm{d}\eta\mathrm{e}^{-2\pi ik\zeta}\,\mathrm{d}\zeta, \end{aligned} \tag{43}$$

where $\widehat{\Theta}_b = \mathrm{e}^{-2\pi|k|y_2}\int \theta_b \mathrm{e}^{-2\pi k\zeta}\,\mathrm{d}\zeta = \mathrm{e}^{-2\pi|k|y_2}\hat{\theta}_b$ and $\widehat{\Theta}_c = \mathrm{e}^{-y_2\sqrt{R^{-2}+4\pi^2k^2}}\hat{\theta}_c$. The integrals in the mer-

dional source variable η are straightforward to compute and are simple exponentials involving the southern frontal position, L_-. Expanding them assuming weak frontal displacements eventually leads to

$$\frac{H_1 k\pi}{\sqrt{R^{-2}+4\pi^2k^2}}\hat{\Theta}_c - \frac{\mathrm{sign}(k)}{2}\hat{\Theta}_b = \left[\frac{-H_1 f_0 q_0}{2}\mathrm{sign}(k) + \frac{f_0 q_0 H_1 \pi k}{\sqrt{R^{-2}+4\pi^2k^2}}\right]\hat{L}_- \quad (44)$$

Comparable manipulations of the other constraints ultimately returns three linearly coupled equations for the Fourier transforms of the unknown potential vorticity fields Θ_1, Θ_b and Θ_c. In all cases, exponentials of the northern and southern frontal locations are involved, which have the effect of making the solution independent of the arbitrary latitudes y_1 and y_2. Solving the system yields explicit expressions for the artificial potential vorticity fields. The most useful result is for the augmentation to the northern, barotropic layer potential vorticity, i.e.

$$\hat{\Theta}_1 e^{2\pi|k|y_1} = \hat{\Theta}_1 = \frac{f_0 q_0 H_2\left(1-\frac{2\pi|k|}{\sqrt{R^{-2}+4\pi^2k^2}}\right)(\hat{L}_+ - \pi|k|\widehat{L_+^2})}{\frac{2(H+H_1)\pi|k|}{\sqrt{R^{-2}+4\pi^2k^2}}+H_2} - q_0 f_0 \times \frac{\frac{4H\pi|k|}{\sqrt{R^{-2}+4\pi^2k^2}}\hat{L}_- + \left(H_2 2\pi|k| + H_1\frac{4\pi^2k^2}{\sqrt{R^{-2}+4\pi^2k^2}}\right)\widehat{L_-^2}}{\frac{(H+H_1)2\pi|k|}{\sqrt{R^{-2}+4\pi^2k^2}}+H_2} \quad (45)$$

because at small amplitude (42) requires only information at the break $y=0$ and can be evaluated from p_1 north of the break alone.

References

Cummins, P., Mysak, L., 1988. A quasi-geostrophic circulation model of the North East Pacific. Part I. A preliminary numerical experiment. Journal of Physical Oceanography 18, 1261–1286.

Hogg, N., 1981. Topographic waves along 70W on the continental rise. Journal of Marine Research 39, 627–649.

Longuet-Higgins, M., 1968. On the trapping of waves along a discontinuity of depth in a rotating ocean. Journal of Fluid Mechanics 31, 417–434.

Mysak, L., Johnson, E., Hsieh, W., 1981. Baroclinic and barotropic instabilities of coastal currents. Journal of Physical Oceanography 11, 209–230.

Pedlosky, J., 2000. The transmission and transformation of baroclinic Rossby waves by topography. Journal of Physical Oceanography 30, 3077–3101.

Pickart, R., 1995. Gulf Stream-generated topographic Rossby waves. Journal of Physical Oceanography 25, 574–586.

Pickart, R., Watts, D., 1990. Deep western boundary current variability at Cape Hatteras. Journal of Marine Research 48, 765–791.

Pratt, L., Stern, M., 1986. Dynamics of potential vorticity fronts and eddy detachment. Journal of Physical Oceanography 16, 1101–1120.

Rhines, P., 1969. Slow oscillations in an ocean of varying depth, Parts I and II. Journal of Fluid Mechanics 37, 161–205.

Rhines, P., 1970. Edge- bottom and Rossby waves in a rotating stratified fluid. Geophysics and Fluid Dynamics 1, 273–302.

Rhines, P., 1977. The Dynamics of Unsteady Currents. In: Goldberg, E. (Ed.), The Sea: Ideas and Observations on Progress in the Study of the Seas. Wiley, New York.

Schultz, J., 1987. Structure and propagation of topographic Rossby waves northeast of Cape Hatteras, North Carolina, M.S. Thesis, Marine Sciences Program, University of North Carolina, 63pp.

Thompson, R., Luyten, J., 1976. Evidence for bottom-trapped topographic Rossby waves from single moorings. Deep-Sea Research 23, 629–635.

Willmott, A., 1984. Forced double Kelvin waves in a rotating, stratified basin. Journal of Marine Research 42, 319–358.

Willmott, A., Johnson, E., 1989. Topographic waves in a rotating statified basin. Geophysical and Astrophysical Fluid Dynamics 45, 71–87.

Willmott, A., Johnson, E., 1995. On geostrophic adjustment of a two-layer, uniformly rotating fluid in the presence of a step escarpment. Journal of Marine Research 53, 49–77.

LSEVIER

Deep-Sea Research II 51 (2004) 2995–3008

DEEP-SEA RESEARCH
PART II

www.elsevier.com/locate/dsr2

Internal tides in the Bay of Biscay: conversion rates and seasonal effects

Theo Gerkema[a,*,1], Frans-Peter A. Lam[b,2], Leo R.M. Maas[b]

[a]*Laboratoire des Ecoulements Géophysiques et Industriels (LEGI), BP 53, 38041 Grenoble Cedex 9, France*
[b]*Royal Netherlands Institute for Sea Research (NIOZ), P.O. Box 59, 1790 AB Texel, The Netherlands*

bstract

Internal tides in the Bay of Biscay are studied using a numerical linear hydrostatic internal-tide generation model, in hich along-slope uniformity is assumed. The focus is on the seasonal changes in the generation and dynamics. The esults are compared with observations made near the shelf-break in early summer. They were obtained using a towed coustic Doppler Current Profiler (ADCP) on repeated tracks in the Bay of Biscay, this provides a synoptic view in a lane spanned by the vertical and the cross-slope direction. In the reproduction of the observed patterns by the umerical model, the presence of the seasonal thermocline turns out to be essential. However, the model results also ndicate that its presence has only a minor effect on the integrated conversion rate (which is about 10 kW/m). The main egion of generation lies deep and is seasonally independent; the presence of the deep permanent pycnocline does ncrease the integrated conversion rate strongly, by a factor two.
© 2004 Elsevier Ltd. All rights reserved.

. Introduction

The Bay of Biscay is thought to offer one of the world's strongest generation sites for internal tides. For this region, Baines (1982) obtained a relatively high estimate of the conversion rate (from barotropic into baroclinic tides), and Pingree and New (1991) observed a distinct internal-tide beam, in which vertical displacements were as large as 300 m (peak-to-trough).

Baines' analysis yielded estimates of conversion rates along continental slopes worldwide; they were later complemented by Morozov (1995) for deep-ocean ridges. Munk (1997) argued that the two-dimensionality assumed in the analysis of Baines (1982) is too restrictive, and that three-dimensional features like canyons are likely to contribute significantly. This presumes that the along-slope barotropic tidal flux is much stronger than the across-slope one, which is however not

*Corresponding author.
E-mail address: gerk@nioz.nl (T. Gerkema).
[1]Present affiliation: Royal Netherlands Institute for Sea Research (NIOZ), P.O. Box 59, 1790 AB Texel, The Netherlands.
[2]Present affiliation: TNO Physics and Electronics Laboratory (TNO-FEL), Underwater Acoustics Department, P.O. Box 6864, 2509 JG The Hague, The Netherlands.

967-0645/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
oi:10.1016/j.dsr2.2004.09.012

the case in the Bay of Biscay (LeCann, 1990; Lam et al., 2004). Yet, even in a purely two-dimensional setting there is ample room for reconsidering Baines' estimates. For example, Baines (1982) adopted a linear slope; as a consequence, the forcing is concentrated at the shelf-break, and the conversion rate shows a strong seasonal dependence. For a more realistic topography, the region of strong forcing (conversion) will be extended, or even shifted, to deeper positions. One must expect that the effect of seasonality loses importance as the deeper (and seasonally independent) generation gains importance. We will argue below that Baines (1982) probably overestimated the conversion rate in the seasonal thermocline, while underestimating the conversion rate in the deep continuously stratified layer. A correction of the latter diminishes the importance of seasonality, and increases the overall conversion rate.

Baines (1982) included seasonal effects by separating the interfacial tide from the tidal beam. The interfacial tide is associated with the seasonal thermocline, which he represented by a δ-function in N^2 (N the buoyancy frequency); the tidal beam is associated with the much weaker deep stratification, which he represented by a layer of constant N. As Baines points out, such a separation is possible if the thermocline is sufficiently strong, because in that case a distinct interfacial mode exists (the first mode), while all higher modes combine to form the beam below the thermocline. It was suggested in Gerkema (2001) that the thermocline in the Bay of Biscay may in fact not be sufficiently strong, implying that no distinct interfacial mode exists, and resulting in a hybrid beam-interfacial manifestation of the internal tide. This will be confirmed below.

Looked upon in terms of ray-theory, instead of a modal decomposition, this hybrid behaviour can be interpreted as a scattering of the beam, the second theme of this paper. Some explanation of this usage of the word 'scattering' may be appropriate. Internal tides, and internal waves in general, owe their existence to the stratification, a measure of which is N. Variations in $N(z)$ affect the propagation of internal waves in two ways. First, it makes the direction of energy propagation vary (refraction), according to $\tan^2\theta = (N^2(z) - \sigma^2)/(\sigma^2 - f^2)$, where θ denotes the angle between the direction of energy-propagation and the vertical z. Here, σ is the tidal frequency, and f the Coriolis parameter. As long as N varies weakly on the typical vertical scale of the internal wave this will be the sole effect. For more strongly varying N, a second effect comes into play: internal reflections, also called secondary reflections (Brekhovskikh, 1960). Multiple secondary reflections cause an initially confined beam to spread: its energy gets scattered over a wide area, but with relatively much energy staying in the thermocline and the mixed layer (as we shall illustrate below). For a very strong thin thermocline, finally, the beam would simply reflect at it, like at a rigid surface, and no significant scattering would occur (this is the regime in which a distinct interfacial mode exists). These processes were studied analytically in a simple setting in Gerkema (2001), where the stratification model of Baines (1982) was adopted. In this paper, the occurrence of scattering will be studied for a more realistic profile of N and topography, by using a numerical model. Since the scattering occurs predominantly in the seasonal thermocline, the seasonal character of scattering is clear a priori, and will indeed be confirmed in this paper.

The outline is as follows. First we present ADCP-observations obtained near the shelf-break during early summer (Section 2.1). The results are compared with those of an internal-tide generation model; model results for winter stratification are also presented (Section 2.2). Next we consider the model results concerning energy-density and conversion rates in the shelf-break region, and compare summer and winter conditions (Section 3). In this section we also reconsider the earlier estimates obtained by Baines (1982). Finally, we show the model results for the far-field (the central Bay of Biscay), and discuss the effects of scattering of a beam at the seasonal thermocline (Section 4).

2. The generation region—current fields

To check whether the model yields realistic results, we compare its outcome with observations

n internal-tide currents, obtained near the shelf-reak, which are briefly described below.

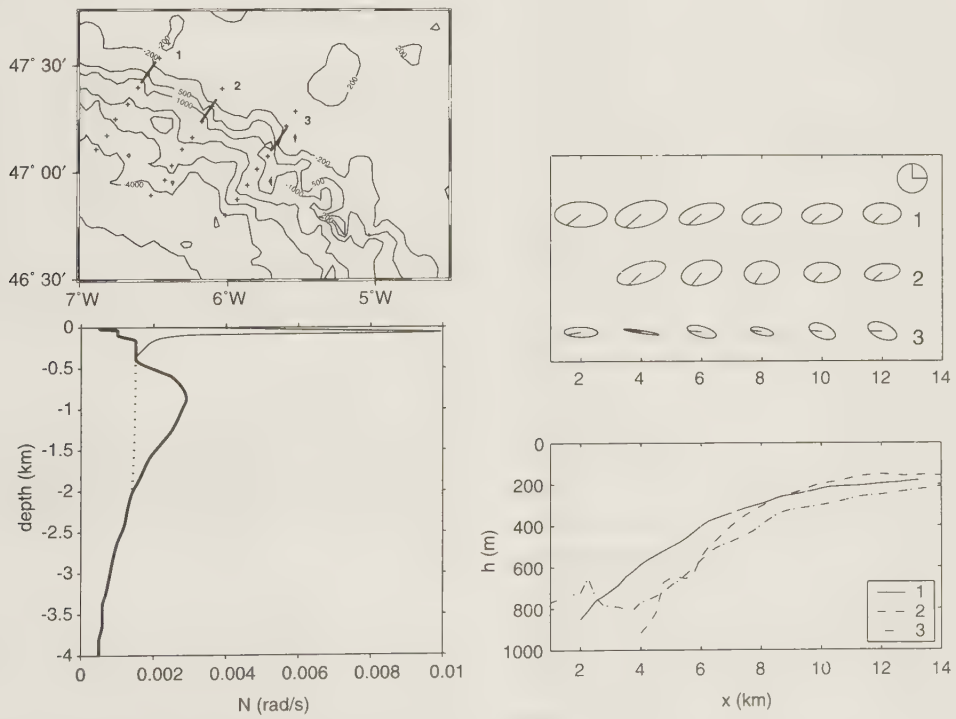
1. Observations

From 4 to 8 June 1993 three cross-isobath ransects, about 35 km apart and 13 km long, were nonitored over the shelf-break in the Bay of iscay (see Fig. 1), each for about 1 day, as part of ne *Triple B* '93 project (Van Aken, 1995). A ownward-looking narrowband 75 kHz ADCP as towed by the ship, measuring the three-imensional currents from 10 m below the water urface to the sea floor in 'bins' with a thickness of m. The tracks were repeated, and thus each osition on the track was visited at least 10 times. his allows one to fit a sinusoidal tidal signal to ach bin; in the fit, an additive constant represents the residual signal. More details on the observational techniques and on the analysis of the signal, including results on the barotropic tidal currents, are presented elsewhere (Lam et al., 2004). Here we focus on the baroclinic tidal signal, which we obtain (approximately) by subtracting the depth-averaged signal; examples of depth-averaged tidal ellipses are shown in Fig. 1.

At each position the fitted baroclinic tidal signal of the cross-slope component can be written as

$$u = A(x,z)\sin[\sigma t - \phi(x,z)], \tag{1}$$

thus lines of equal phase propagate in the direction of increasing ϕ. The resulting amplitude A and phase ϕ for one transect (the other two are similar) are shown in Fig. 2. A number of conspicuous features is revealed in these observations. First of all, the distinctive blue band in Fig. 2A, where



ig. 1. Left-hand panels: bathymetry and transects where ADCP measurements were made (the results of transect 1 are used in this aper); and (lower panel) the stratification profiles used in the model calculations: summer (thin line), winter (thick line), and a ctitious 'winter' profile in which the permanent pycnocline has been removed (dots). Right-hand panels: barotropic tidal ellipses (in erms of transport) over each of the transects (topography is shown in the lower panel); the horizontal axis of the ellipse gives the cross-ope transport, the vertical axis the along-slope transport; notice that the former exceeds the latter in all cases. The radius in the eference circle indicates a flux of $100\,\mathrm{m^2\,s^{-1}}$.

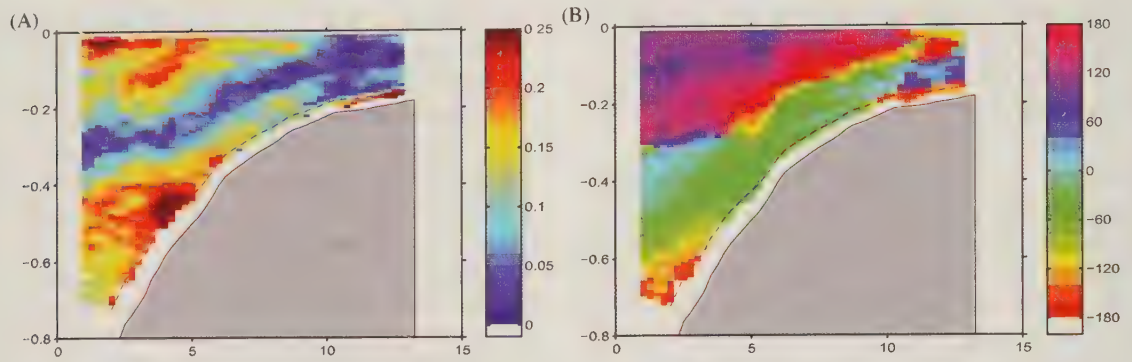


Fig. 2. Results from ADCP observations in the Bay of Biscay ($x = 0$: 47.42°N; 6.59°W): the baroclinic cross-slope tidal current, decomposed in (A) its amplitude (in $\mathrm{m\,s^{-1}}$) and (B) its phase (in degrees). Distances along the axes are in km. The x-axis is directed 34.2° True North.

currents are zero, or nearly so; it runs from the middle left to the upper right, and divides the figure into two parts. In this band, an amphidromic point occurs near $x = 4.5$ km, $z = -270$ m. By an amphidromic point we understand a position at which the currents vanish while co-phase lines circle around it. In the context of internal tides, this phenomenon was first identified by DeWitt et al. (1986). To the left of this point, phase propagation is generally upward (except in the upper 150 m); to the right, it is generally downward (Fig. 2B).

In particular, in the region at the lower left (i.e. below the blue band), where currents are particularly strong, phase propagation is upward. Basic results from internal-wave theory indicate that the group-velocity (i.e. the direction of energy-propagation) is perpendicular to the direction of phase-propagation (i.e. parallel to co-phase lines), with opposing vertical components (LeBlond and Mysak, 1978). Applying these theoretical results, we find that internal-tide energy propagates downward in this region, parallel to the co-phase lines. As we shall discuss below, this signal in fact belongs to the most upper part of an internal-tide beam similar to the one observed by Pingree and New (1991), further off the shelf-break.

The region in the upper left corner (i.e. above the blue band) also shows strong currents, but the direction of phase propagation is less distinctive. With the use of the numerical model, we shall suggest an interpretation of this phenomenon.

2.2. Model results

In this study we use a linear hydrostatic internal-tide generation model. It is a slightly modified version of the model described in Gerkema (2002); a more accurate numerical scheme is now used (see below). The model can be employed for arbitrary stratification $N(z)$ and topography $z = h(x)$ (x is the across-slope direction, z the vertical). Uniformity is assumed in the along-slope direction: $\partial/\partial y = 0$.

2.2.1. Model equations

We use the inviscid linear hydrostatic equations on the f-plane, under the Boussinesq approximation (partial derivatives are denoted by indices):

$$\psi_{zzt} - f v_z + b_x = 0, \tag{2}$$

$$v_t + f\psi_z = 0, \tag{3}$$

$$b_t - N^2\psi_x = -N^2 W. \tag{4}$$

Here ψ is the streamfunction ($u = \psi_z, w = -\psi_x$); v the along-slope velocity component; b denotes buoyancy, defined as 'minus effective gravity': $b = -g\rho/\rho_*$, where ρ is the density's local departure from its hydrostatic value, and ρ_* a constant reference value of density. Isopycnal displacements (ζ) can be derived from buoyancy via $b = -N^2\zeta$.

For the stream function we impose the boundary conditions $\psi = 0$ at bottom and surface (rigid-lid).

The barotropic forcing is prescribed via W, the vertical velocity component of the barotropic tide, an expression of which is obtained in the following way. We assume that the cross-slope barotropic tidal flow U is such that it produces a spatially constant but time-oscillating flux: $U(t,x)h(x) = Q_0 \sin \sigma t$ (constant Q_0). Via continuity, $U_x + W_z = 0$, we then find for W:

$$W = z \frac{Q_0 \sin \sigma t}{h(x)^2} h_x , \qquad (5)$$

where the integration constant was chosen such that $W|_{z=0} = 0$.

The model equations do not only produce time-oscillating fields, but also mean fields for b and v, depending on the initial conditions. This can be seen as follows: (3) and (4) imply

$$v_x + f\left(\frac{b}{N^2}\right)_z = f \frac{Q_0 \cos \sigma t}{\sigma h^2} h_x + c(x,z) , \qquad (6)$$

where $c(x,z)$ is a 'constant' of integration due to time-integration. This constant will be nonzero if we choose (as is natural) $v = b = 0$ at $t = 0$; as a consequence, v and b will then contain stationary fields at later times. These fields form an arbitrary and artificial element in the solution, since they depend on how we choose the initial conditions. They reflect the 'memory' of the system over the topography due to Coriolis effects, and (6) can in fact be interpreted as an expression of conservation of potential vorticity. As expected, we found that the stationary fields become negligible if we replace $\sin \sigma t$ in (5) by $\cos \sigma t$ (giving $\sin \sigma t$ in (6)), but this abrupt way of starting the forcing will usually be less attractive from a numerical point of view.

2.2.2. Numerical scheme and resolution

A new vertical coordinate is introduced: $\eta = 1 - 2z/h(x)$, which transforms the vertical domain to $(-1,1)$. The Eqs. (2)–(5) are transformed accordingly. The boundary conditions then become $\psi = 0$ at $\eta = \pm 1$. In the vertical, a Chebyshev collocation method is used, the collocation points being defined by $\eta_k = \cos(\pi(k-1)/(K-1))$ for $k = 1, \ldots, K$. This gives a higher resolution near the surface (and bottom), which improves the resolution in the seasonal thermocline and the description of the reflection of beams at surface and bottom.

We use central differences in x (5-point symmetric, 4th-order), and perform the time-integration using a third-order Adams–Bashforth scheme. A fourth-order spatial filter is used to selectively dampen oscillations on the grid scale (Durran, 1999). Sponge layers are used by including Rayleigh friction, the frictional coefficient being zero at the entrance of the sponge, and increasing linearly inward. The thickness of the sponge is 150 km in the deep ocean, and 50 km on the shelf.

In all calculations shown below (except those in Section 3.1) the horizontal step is 400 m; this means that the continental slope is covered by some 100 points. In the vertical a sufficiently good resolution is obtained by using 80 collocation points. The number of time steps per tidal period (M2) is 800, giving a time step of nearly 56 s. Each run spans 60 tidal periods; this length is sufficient to ensure that all transients have left the domain under consideration by the end of the run. The results we show are taken from the last period.

2.2.3. Results

For the stratification we use an observed profile (Fig. 1, thin line), in which a pronounced seasonal thermocline is present. We also do calculations for a winter profile, in which the thermocline has been removed (thick line). In both profiles a clear permanent pycnocline is present (around 1 km depth), which is typical of the Bay of Biscay (Van Aken, 2001). In one model calculation (next section), we use a fictitious profile in which the permanent pycnocline has been removed as well (dotted line in Fig. 1). For the topography we use data from an observed nearby transect traversed in a later cruise, which covered the whole slope. For Q_0, the amplitude of the barotropic cross-slope flux, we choose $100\,\mathrm{m^2\,s^{-1}}$, being a typical value for this region (see Fig. 1).

We present the results in terms of the cross-slope baroclinic current u, decomposed into amplitude and phase fields as in (1), see Fig. 3. To facilitate the interpretation, a larger domain is shown than in the observations; the observational domain is

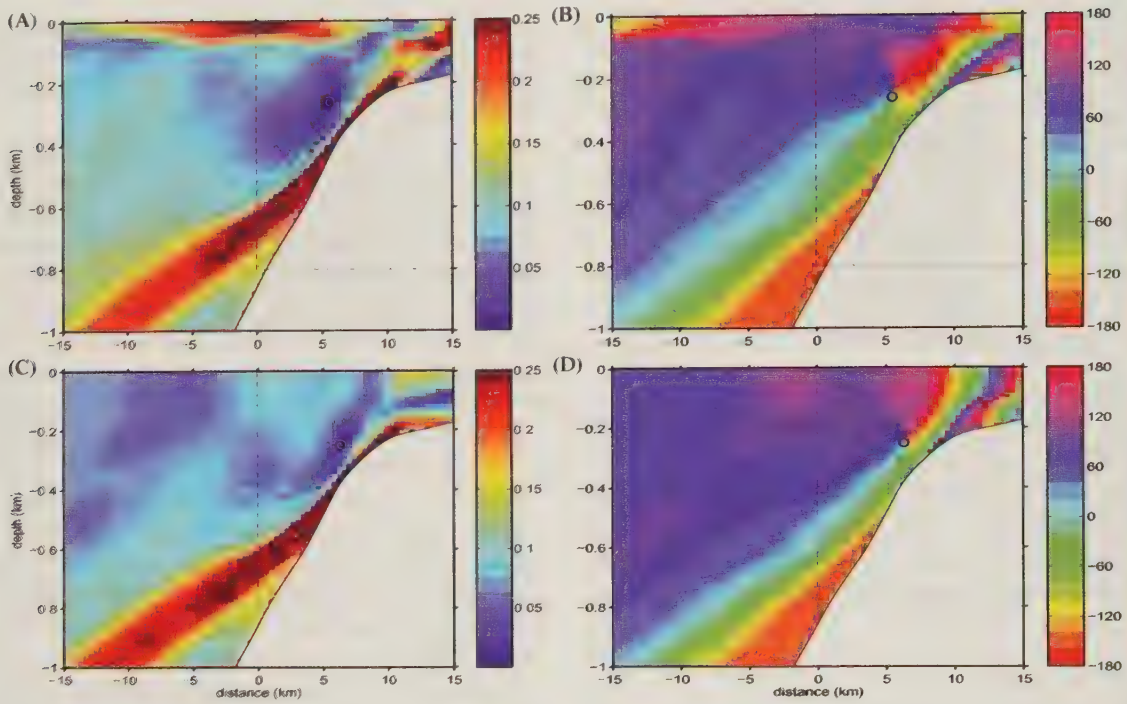


Fig. 3. Model results: the baroclinic cross-slope component ($u = \psi_z$), decomposed into: (A, C) amplitude A (in m s^{-1}); (B, D) phase ϕ (in degrees), for two different stratifications: (A, B) summer; (C, D) winter. The box delineates the region of observation, see Fig. 2. The circles indicate the presence of an amphidromic point.

indicated by a box. In Figs. 3A and B the results are shown for the observed (summer) stratification. The model reproduces the diagonal blue band of near-zero currents, the small region of strong currents in the upper-left part, and the strong currents in the lower-left part, initially attached to the topography. Also the amphidromic point is well reproduced, although not exactly at the same location: $x = 5.6$ km, $z = -250$ m (see circles), which is more to the right, and slightly higher, than in the observations. The strong tidal currents in the lower-left part are (as in the observations) associated with upward phase propagation, hence downward energy propagation along the co-phase lines; the beam-like nature is clearly seen in Fig. 3A. In the upper-left part, the direction of phase propagation is unclear, like in the observations.

Finally, at the right-end of the domain, the internal tides propagate on-shelf. There the pattern corresponds less well with that of the observations. This might be due to barotropic advection (neglected in the model), which can be expected to become significant in shallow regions in that it affects the propagation of the internal tide (modifying its effective phase speed), as was observed by Pingree et al. (1986). Another cause of the difference may lie in horizontal changes in stratification on the shelf-side. Since the focus of this paper is on the internal tides travelling into the deep ocean, we shall not pursue this problem here.

For completeness, we also present the results for winter statification (no seasonal thermocline, see Fig. 1), all other parameters being the same. In this case, we have no observations to compare with. In the upper 200 m, the pattern (see Figs. 3C and D) looks very different from that during summer. Only the strong currents in the deep lower-left part have 'survived' the removal of the seasonal thermocline, and form (as in Fig. 3A) the beginning of a distinct beam that propagates into the abyssal ocean. In the rest of this paper, we shall

efer to this beam as the *main beam*; it is equally resent in summer and winter.

The generation region—conversion rates

From (2) to (4) one can derive the following nergy equation:

$$E_t + \rho_*[f(\psi v)_z - (\psi\psi_{zt})_z - (\psi b)_x] = -\rho_* b\, W\,, \quad (7)$$

where E denotes the energy-density:

$$E = \frac{1}{2}\,\rho_*\{(\psi_z)^2 + v^2 + b^2/N^2\}\,. \quad (8)$$

The term $\rho_*[\cdots]$ on the left-hand side of (7) can lso be written in the familiar flux-form $\nabla\cdot(p\vec{u})$, where p is the baroclinic pressure. For a purely ime-periodic signal, the first term on the left-hand ide (7) yields zero if we integrate over one tidal eriod T, hence

$$\rho_*\langle f(\psi v)_z - (\psi\psi_{zt})_z - (\psi b)_x\rangle = -\rho_*\langle b\, W\rangle\,, \quad (9)$$

where $\langle\,\rangle = T^{-1}\int_0^T \mathrm{d}t$. We define

$$C = -\rho_*\langle b\, W\rangle \quad (10)$$

s the local *conversion rate*; C indicates how much nergy per second per unit volume is converted rom barotropic into baroclinic tides. Notice that ne cannot calculate C a priori, because it contains he baroclinic field b, which first has to be solved rom the model equations. This reflects the fact hat the amount of energy converted at some ocations depends not only on the local forcing W) but also on the signal produced elsewhere.

The isopycnal displacement ζ (and also the uoyancy, b) includes the purely barotropically nduced vertical elevations (over the topography). hese elevations correctly give no contribution to he right-hand side of (10), because the corre- ponding b and W are 90° out of phase. The rtificial mean fields discussed in Section 2.2.1 give o contribution, either. This can be seen by writing $= \tilde{b}(x,z) + b'(t,x,z)$, where $\tilde{b}$ denotes the mean eld, while b' is purely oscillatory. Now, $\tilde{b}$ bviously leaves the conversion rate C unaffected, nd it does not modify b' either (which would ndirectly affect C), since the model is linear. Thus, ne arbitrariness in the solution does, fortunately, ot extend to the important quantity C. However, the mean fields $\tilde{b}$ can be expected to disturb C in nonlinear models, for we found that $\tilde{b}$ is relatively large in the main region of generation; the outcome of conversion rates in nonlinear models should therefore be regarded with caution.

In view of the two-dimensionality of our problem, we define the 'integrated conversion rate' C_{int} as the integral of C over the xz-plane, being equivalent to the 'total energy flux' in Baines (1982). The equivalence becomes clear if we integrate (9) over the horizontal domain (x_1, x_2), covering the entire generation region, and over the vertical domain $(h(x), 0)$:

$$-\rho_*\left(\int \mathrm{d}z\ \langle\psi b\rangle\right)\Big|_{x_1}^{x_2} = \int \mathrm{d}x \int \mathrm{d}z\ C \equiv C_{\mathrm{int}}\,. \quad (11)$$

The expression on the left-hand side is also useful in determining the integrated conversion rate empirically, since $b = -N^2\zeta$ and ψ are relatively easy to obtain.

Before showing the results for C and C_{int} for the region of observation (Section 3.2), we first consider the idealized case of a linear slope to place our estimates in the context of those obtained before.

3.1. Linear slopes: comparison with earlier estimates

Recent analytical studies on the conversion rate were made by Llewellynn Smith and Young (2002) and Khatiwala (2003), under the assumption of infinitesimal topography. (Notice that Llewellynn Smith and Young (2002) understand by the 'conversion rate' the quantity C as above, but integrated over all three spatial directions.) For continental slopes, i.e. finite-amplitude topography, an analytical result is known for the special case of a step-topography. A solution was derived by Sjöberg and Stigebrandt (1992), who however neglected the on-shelf propagating internal tide. Their solution has recently been generalized to include the on-shelf contribution as well (St. Laurent et al., 2003; Sanchis, 2003).

For linear slopes other than a step no closed expression has been obtained so far. In a semi-analytical approach, Baines (1982) obtained

estimates for the conversion rates in the seasonal thermocline (interfacial tides) and in the constantly stratified deep layer beneath it (internal-tide beams). For the Bay of Biscay, he found 0.936 and $0.232\,\mathrm{kW\,m^{-1}}$, respectively, implying that seasonality is a dominant factor in the conversion rates. However, the latter value deserves further scrutiny; we shall reconsider it in isolation by momentarily ignoring the seasonal thermocline.

We have done a series of calculations, using the following parameters: latitude 45°N, constant buoyancy frequency $N = 5.5 \times 10^{-3}\,\mathrm{rad\,s^{-1}}$, shelf depth 135 m, ocean depth 4 km, M2 tidal-frequency $\sigma = 1.405 \times 10^{-4}\,\mathrm{rad\,s^{-1}}$, $\rho_* = 1025\,\mathrm{kg\,m^{-3}}$. These values correspond essentially to those used by Baines (1982), except that we have ignored the mixed layer of 100 m thickness, which he included by taking $N = 0$ there (see below). Two crucially important parameters, however, were not specified: the cross-slope barotropic flux Q_0 and the slope length L. For the former we take $100\,\mathrm{m^2\,s^{-1}}$, and we will vary L from 10 to 50 km. In treating the generation at a linear slope numerically, one does not easily find accurate values of the integrated conversion rate (and proper convergence for finer resolution), because the conversion is largely concentrated at the sharp upper corner of the slope. We have therefore replaced the corner, over an interval 2ε, by a parabolic profile such that h becomes continuously differentiable (we have similarly replaced the corner at the foot of the slope). This is a mathematically consistent way of smoothing, since the second derivative of h is now piecewise constant and properly yields two δ-distributions in the limit $\varepsilon \to 0$. We chose $\varepsilon = 2$ km; we used 100 collocation points in the vertical, a horizontal step of 100 m, and 4000 time steps per tidal period (six tidal periods in total).

The results are gathered in Fig. 4; they allow us to draw three conclusions. First, in this regime of steep slopes, the value of C_{int} varies little with slope length L, and is always close to the theoretical value for a step (St. Laurent et al., 2003), indicated by the horizontal dotted line ($36.6\,\mathrm{kW\,m^{-1}}$). Second, the values for C_{int} in Fig. 4 are *two* orders of magnitude larger than Baines' (which was $0.232\,\mathrm{kW\,m^{-1}}$). Third, the preceding two points imply that the discrepancy cannot be

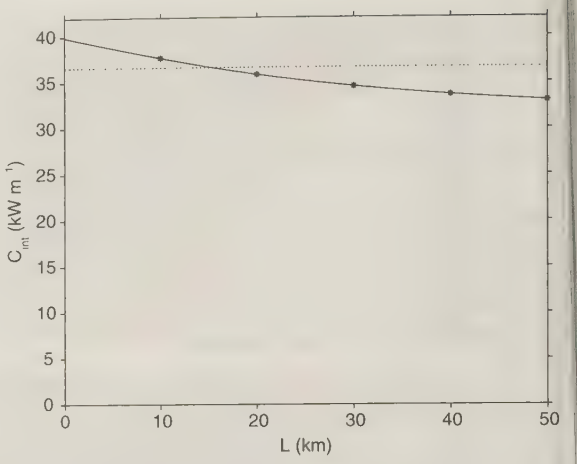


Fig. 4. The integrated conversion rate C_{int} as a function of slope length L, for near-linear slopes (points); the solid line results from a cubic interpolation (and extrapolation for $L < 10$ km), and the horizontal dotted line represents the theoretical exact value for a step-topography.

ascribed to a possible difference in the value of L. Furthermore, we found that adding a mixed layer of 100 m thickness (as had Baines) reduces C_{int} only by about 2%.

If we assume that Baines' estimate is technically accurate (we notice, however, that his derivation involves a few approximations), then the explanation of the discrepancy must stem from a difference in the choice of the barotropic cross-slope flux, Q_0. Our value of $100\,\mathrm{m^2\,s^{-1}}$, although appropriate for the region we consider (see Fig. 1), is indeed not representative for the Bay of Biscay as a whole. However, to arrive at Baines' value for C_{int}, we would have to reduce Q_0 by a factor of ten (since C_{int} is proportional to Q_0^2), giving an unrealistically small Q_0. Judging from observations and calculations made by LeCann (1990), we may conclude that a representative value is $Q_0 = 30 - 40\,\mathrm{m^2\,s^{-1}}$; this would reduce the value for C_{int} from Fig. 4 by one order of magnitude, leaving it still *one* order of magnitude higher than Baines'. (As an aside, we notice that Fig. 7B in LeCann (1990), like our Fig. 1, demonstrates clearly that cross-slope fluxes are generally larger than along-slope fluxes.)

Although questions remain, it seems safe to conclude that Baines' estimate of the contribution

of the continuously stratified deep layer to C_{int} is too small by one order of magnitude.

3.2. Results

We now return to the model results for the Bay of Biscay for the observed topography and stratification. In Fig. 5 we show the tidal-averaged energy density $\langle E \rangle$, and the conversion rate C, both for summer and winter. C_{int}, the integrated value of C, is also indicated in the figures. As mentioned above, the integration is made over the xz-plane, so that the resulting quantity is in Watt per meter, which can be interpreted as Watt per unit slope-length (in the transverse direction, y). During summer, one can identify two well-separated source regions (Fig. 5B): a relatively deep one, attached to the slope, and a small region higher up, which lies in the seasonal thermocline. The latter is indeed absent during winter (Fig. 5D). It is now clear that the seasonal thermocline forms the source-region of the strong internal-tide currents in the upper-layer on the left as well as of the on-shelf propagating internal tides (Fig. 2A, Fig. 3A); both are absent during winter (Fig. 3C). Although the baroclinic fields look very different in summer and winter, notably in the upper 200 m (Figs. 5A and C), the integrated conversion rates C_{int} are fairly comparable (Figs. 5B and D); during winter, its value is only 15% smaller. In other words, the model indicates that the manifestation of internal tides is very different in summer and winter, while their total energy is not. Figs. 5B and D show clearly why there is little seasonal dependence in C_{int}: the main region of energy conversion lies well below the seasonal thermocline; it is also the region from which the main beam originates.

The values for C_{int} obtained here lie still remarkably close to the value for the step-topography

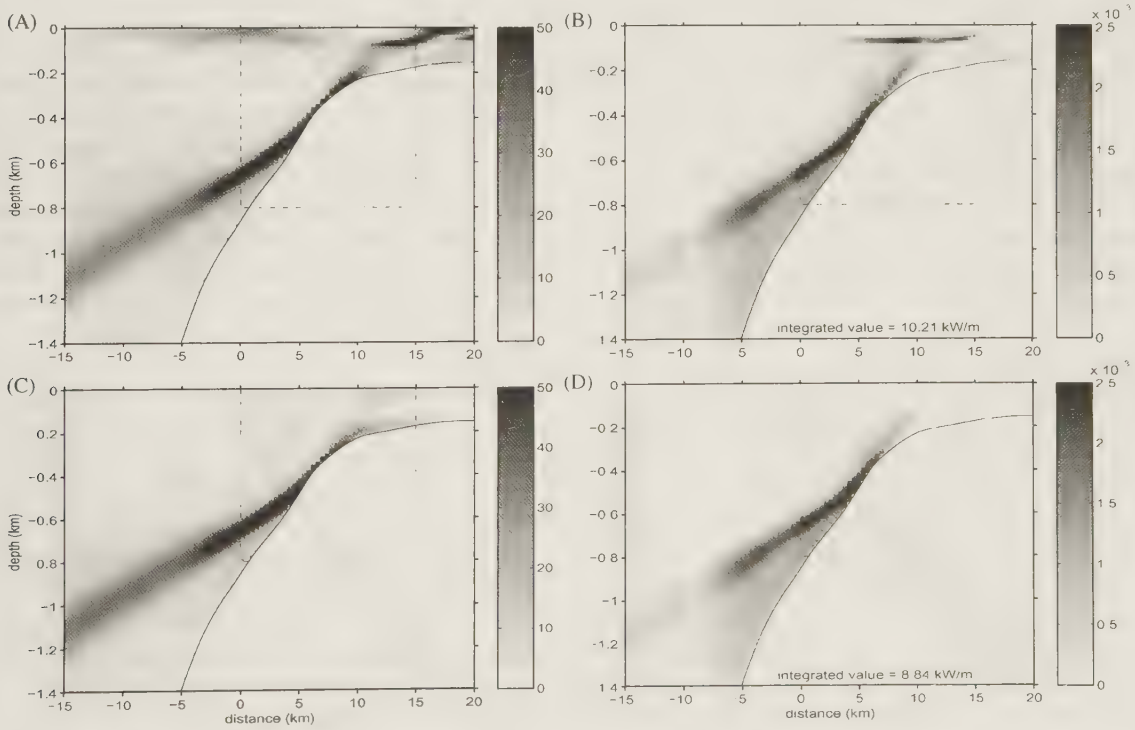


Fig. 5. Model results: (A) the tidal-averaged baroclinic energy-density $\langle E \rangle$ (in J m^{-3}) during summer; (B) the conversion rate C (in W m^{-3}) during summer; (C) $\langle E \rangle$ during winter; (D) C during winter. The box delineates the region of observation. In (B, D) the integrated value over the xz-plane, C_{int}, is also indicated.

and near-linear slopes of Fig. 4, if we allow for the fact that in the main region of generation (see Fig. 5D), N is about two to three times smaller than the value used in Fig. 4. (For a step-topography, C_{int} becomes proportional to N under the hydrostatic approximation, see St. Laurent et al., 2003).

The difference we find between summer and winter (1.36 kW m^{-1}), which can be ascribed to the seasonal thermocline, is of the same order of magnitude as the contribution Baines (1982) found for the seasonal thermocline (0.936 kW m^{-1}). This is remarkable in view of our conclusion, argued above, that Baines probably used a much lower value for the barotropic cross-slope flux, Q_0. However, this will be compensated (at least in part) by the fact that he used a linear slope; the position of maximum W is then located at the upper corner of the slope, which lies close to the seasonal thermocline. For a realistic topography, W takes its maximum value at a much deeper position.

Finally, we consider a fictitious stratification in which the permanent pycnocline has been removed as well (Fig. 1, dotted line). The results (Fig. 6) are qualitatively similar to those during winter, except that the main (and only) beam is now steeper near 1 km depth, due to the decreased value of N. However, quantitatively there is a very significant difference: the integrated conversion rate has been *halved*, compared to the winter situation. This means that the permanent pycnocline is of primary importance for a correct modeling of the energetics and conversion rates.

4. The far field—scattering

We now consider the development of the internal tides as they propagate further into the deep ocean. To this end, we show the baroclinic cross-slope component u (again decomposed in amplitude and phase fields as in (1)), both for summer and winter, see Fig. 7.

The interpretation of the results during winter is most straightforward (Figs. 7C and D): here the main beam is the only beam present. It originates from the upper parts of the slope (cf. Figs. 5C and D), and first moves downward, then reflects at the bottom ($x = -55$ km), and later at the surface ($x = -120$ km). The beam is less intense (less narrow) after its reflection at the bottom than before; this is because the bottom is sloping, and hence defocuses the beam. The direction of the beam changes slightly with the vertical (steeper at deeper positions), which is due to the varying stratification; this is the refraction discussed in Section 1. Essentially, it amounts to a WKB-like adaptation to a weakly varying background field, $N(z)$; no significant scattering occurs. In Fig. 7C we have added the positions at which Pingree and New (1991) found the maximum vertical displacements (see their Fig. 9). Although their

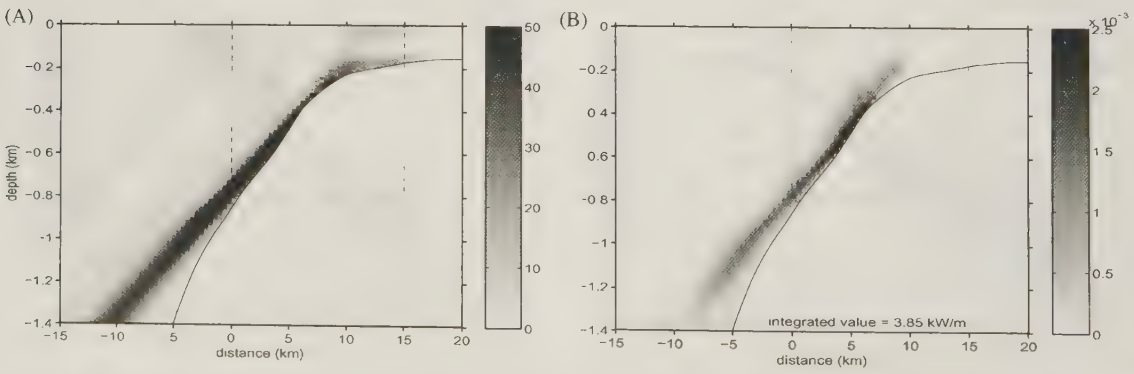


Fig. 6. Model results for a fictitious profile of N, in which the permanent pycnocline has been removed: (A) the tidal-averaged baroclinic energy-density $\langle E \rangle$ (in J m^{-3}); (B) the conversion rate C (in W m^{-3}). The box delineates the region of observation. In (B) the integrated value over the xz-plane, C_{int}, is also indicated.

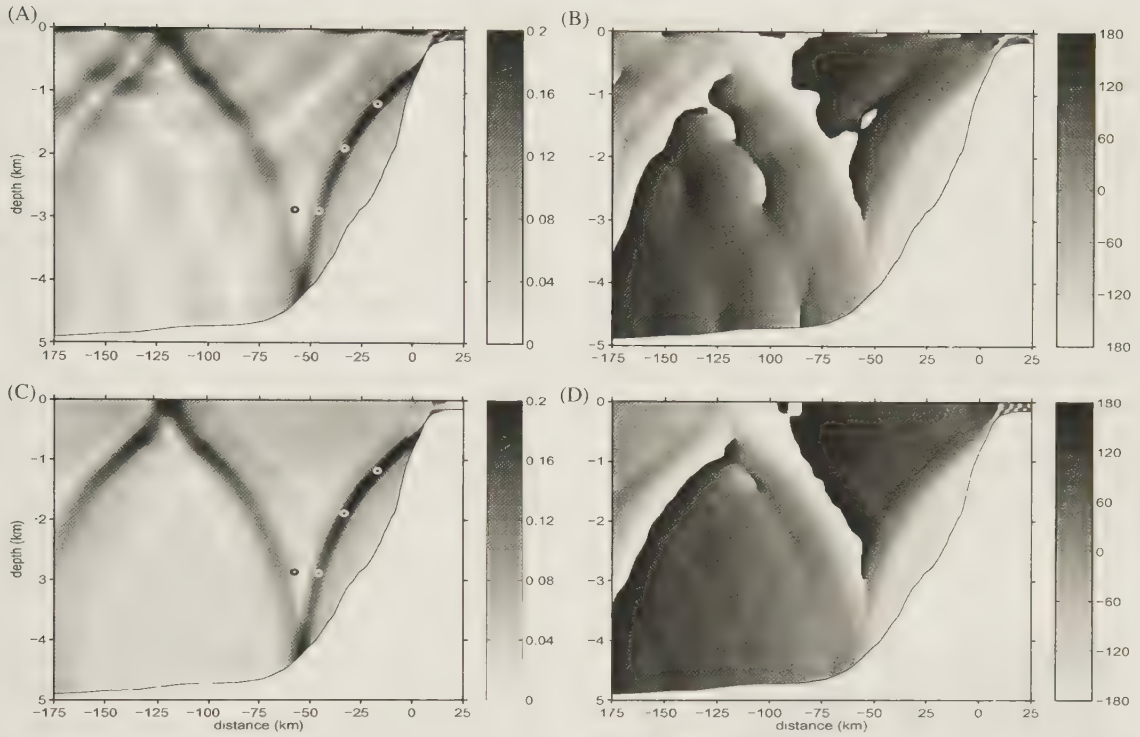


Fig. 7. Model results, large domain: the baroclinic cross-slope component ($u = \psi_z$), decomposed into: (A, C) amplitude A (in m s^{-1}); and (B, D) phase ϕ (in degrees; notice that the extremes of black and white denote the same phase), for two different stratifications: (A, B) summer; (C, D) winter. The circles indicate the location of the main beam as observed by Pingree and New (1991).

measurements were made during summer, one expects no seasonal effects in the beam's initial downward propagation, given the depth of the generation region; the correspondence is indeed very satisfactory.

During summer, the manifestation of internal tides is very different: see Figs. 7A and B. The only part that remains the same as during winter is the initial downward path of the main beam. Above this beam, and parallel to it, one now discerns weak beams (in grey), which originate from the regions in the thermocline/mixed-layer where currents are strong: i.e. near $x = 0, -35$ and -70 km. They interfere with the main beam, once the latter has reflected at the bottom and is directed upward (between $x = -120$ and -55 km); in this traject the main beam looks therefore slightly different from the winter case.

The main beam itself encounters the thermocline near $x = -120$ km. In winter, the beam reflects at the surface without undergoing any distortion (Fig. 7C), but in Fig. 7A the pronounced variation in stratification in the seasonal thermocline causes internal (or secondary) reflections, hence scattering and a severe distortion of the beam. In the presence of nonlinear and non-hydrostatic effects (both are however absent in the present model), the local disturbance of the seasonal thermocline provoked by the incoming beam can give rise to the appearance of internal solitary waves (Gerkema, 2001). For the Bay of Biscay, this phenomenon has been reported in the literature (New and Pingree, 1990, 1992). In these studies the connection with the main beam was established: in the central Bay of Biscay internal solitary waves occur precisely where the main beam encounters the thermocline (near

$x = -120$ km in Fig. 7A). Moreover, they are indeed observed only during summer (New and Da Silva, 2002).

In the region of scattering, relatively much energy remains temporarily behind in and above the thermocline, as the spots of strong tidal currents show. This is more clearly shown in an enlargement, see Fig. 8. In the upper 60 m, phase propagation is predominantly horizontal (Fig. 8B); in this respect the behaviour resembles that of a pure interfacial tide in a two-layer system (with thin upper layer d). The phase speed in such a system is given by $c^2 = g'd$; respresentative values are $g' = 0.01\,\mathrm{m\,s^{-2}}$ and $d = 60$ m. One finds a wavelength of about 35 km, which corresponds well to the wavelength in the upper 60 m in Fig. 8B, and is also fairly close to the observed value of 40–45 km (New and Da Silva, 2002). Yet, the internal tide here differs crucially from an interfacial tide: for the latter, one would find a *constant* amplitude A in the upper layer, which obviously is not the case in Fig. 8A. Instead of a constant amplitude, one observes the spots referred to above, formed by an occasional constructive superposition of the modes, in the absence of a distinct interfacial mode.

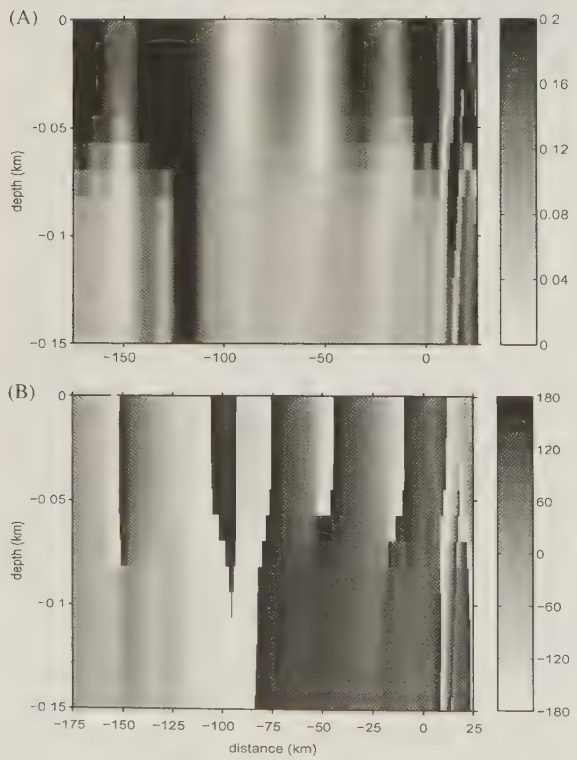


Fig. 8. Model results, upper layer during summer (enlargement of Fig. 7(A, B): the baroclinic cross-slope component, decomposed into: (A) amplitude A (in $\mathrm{m\,s^{-1}}$); and (B) phase ϕ (in degrees; notice that the extremes of black and white denote the same phase).

5. Discussion

The numerical results presented in Section 3.2 show that there are two main well-separated regions of internal-tide generation: one in the seasonal thermocline, and a deeper seasonally independent one (between 300 and 900 m depth). The latter dominates the integrated conversion rate; seasonal influences are, therefore, of secondary importance. This qualitative conclusion can be expected to remain valid if one would include factors ignored here, like three-dimensionality and nonlinearity. The possible importance of three-dimensional features like canyons was stressed by Munk (1997), but for the region considered here they are not likely to change the conversion rate dramatically, because the cross-slope barotropic tidal flux (Q_0) exceeds the along-slope one (see Fig. 1). Nonlinear effects and mixing would locally change the stratification and hence the conversion rates; for example, Xing and Davies (1998) found an increase by a factor three. (The caveat concerning artificial mean fields should be noticed when evaluating conversion rates in the nonlinear problem, see Section 3.)

The integrated conversion rates we found in Section 3.2 (for realistic topography and stratification) exceed those obtained by Baines (1982) by one order of magnitude. Even if we adopt his setting (linear slope and constantly stratified deep layer, Section 3.1), we still find values much higher than his: one order of magnitude higher for the Bay of Biscay as a whole, two orders of magnitude for the location under consideration. (Baines probably used a barotropic cross-slope flux that

is much weaker than the observed value.) In the setting with a linear slope, we verified our model results by comparing them with the exact theoretical result for a step-topography (Fig. 4).

Unfortunately, at this stage it is difficult to verify the conversion rates empirically. Recent developments in the use of Topex/Poseidon satellite observations bear the promise of providing accurate conversion rates sometime in the near future. At present the estimated rates are probably already reasonable in a globally integrated sense (Egbert and Ray, 2000), but still uncertain for individual locations. In situ measurements and an assessment of the energy fluxes (not yet carried out in the Bay of Biscay) would provide a much needed test for theoretical and numerical models.

Although the total energy involved in the internal tides here shows little seasonal dependence, their evolution and propagation changes much with the seasons. First of all, the extra source-region in the seasonal thermocline, though relatively small, gives rise to internal-tide currents in the upper layer at the ocean-side (clearly visible in the observations, too), and to internal tides propagating on-shelf. The dynamics of the main beam, whose generation is not affected by the seasons, shows a marked seasonal dependence too, but not before it approaches the surface in the central part of the Bay. During winter, the beam continues its propagation undisturbed after the reflection at the surface; during summer, it gets severely distorted due to scattering in the seasonal thermocline. As discussed above, this process of scattering can be linked to the occurrence of internal solitary waves.

Acknowledgements

The authors are much indebted to H.M. van Aken, who designed and supervised the field program, and to S. Ober and C. Veth for their assistance with the ADCP observations. T.G. and L.R.M.M. received financial support from the project INTAS01-0025.

References

Baines, P.G., 1982. On internal tide generation models. Deep-Sea Research 29, 307–388.

Brekhovskikh, L.M., 1960. Waves in Layered Media. Academic Press, New York.

DeWitt, L.M., Levine, M.D., Paulson, C.A., Burt, W.V., 1986. Semidiurnal internal tide in JASIN: observations and simulation. Journal of Geophysical Research 91 (C2), 2581 2592.

Durran, D.R., 1999. Numerical Methods for Wave Equations in Geophysical Fluid Dynamics. Springer, Berlin.

Egbert, G.D., Ray, R.D., 2000. Significant dissipation of tidal energy in the deep ocean inferred from satellite altimeter data. Nature 405, 775–778.

Gerkema, T., 2001. Internal and interfacial tides: beam scattering and local generation of solitary waves. Journal of Marine Research 59, 227–255.

Gerkema, T., 2002. Application of an internal-tide generation model to baroclinic spring-neap cycles. Journal of Geophysical Research 107 10.1029/2001JC001177.

Khatiwala, S., 2003. Generation of internal tides in an ocean of finite depth: analytical and numerical calculations. Deep-Sea Research 50 (I), 3–21.

Lam, F.P.A., Maas, L.R.M., Gerkema, T., 2004. Spatial structure of tidal and residual currents as observed over the shelf break in the Bay of Biscay. Deep-Sea Research 51 (8), 1075–1096.

LeBlond, P.H., Mysak, L.A., 1978. Waves in the Ocean. Elsevier Oceanography Series. Elsevier, Amsterdam.

LeCann, B., 1990. Barotropic tidal dynamics of the Bay of Biscay shelf: observations, numerical modelling and physical interpretation. Continental Shelf Research 10 (8), 723–758.

Llewellynn Smith, S.G., Young, W.R., 2002. Conversion of the barotropic tide. Journal of Physical Oceanography 32, 1554–1566.

Morozov, E.G., 1995. Semidiurnal internal wave global field. Deep-Sea Research 42 (1), 135–148.

Munk, W., 1997. Once again: once again—tidal friction. Progress in Oceanography 40 (1–4), 7–35.

New, A.L., Da Silva, J.C.B., 2002. Remote-sensing evidence for the local generation of internal soliton packets in the central Bay of Biscay. Deep-Sea Research 49, 915–934.

New, A.L., Pingree, R.D., 1990. Large-amplitude internal soliton packets in the central Bay of Biscay. Deep-Sea Research 37, 513–524.

New, A.L., Pingree, R.D., 1992. Local generation of internal soliton packets in the central Bay of Biscay. Deep-Sea Research 39 (9), 1521–1534.

Pingree, R.D., Mardell, G.T., New, A.L., 1986. Propagation of internal tides from the upper slopes of the Bay of Biscay. Nature 321, 154–158.

Pingree, R.D., New, A.L., 1991. Abyssal penetration and bottom reflection of internal tidal energy into the Bay of Biscay. Journal of Physical Oceanography 21 (1), 28–39.

Sanchis, A., 2003. Modélisation de la marée interne: méthode analytique et validation numérique. MSc, Thesis, Ecole Centrale de Nantes.

Sjöberg, B., Stigebrandt, A., 1992. Computations of the geographical distribution of the energy flux to mixing processes via internal tides and the associated vertical circulation in the ocean. Deep-Sea Research 39 (2), 269–291.

St. Laurent, L., Stringer, S., Garrett, C., Perrault-Joncas, D., 2003. The generation of internal tides at abrupt topography. Deep-Sea Research 50, 987–1003.

Van Aken, H.M., 1995. Biscay 1993. Cruise Report, Nether lands Institute for Sea Research (NIOZ).

Van Aken, H.M., 2001. The hydrography of the mid latitude. Northeast Atlantic Ocean Part III: The thermocline water masses. Deep-Sea Research 48 (I) 237–267.

Xing, J., Davies, A.M., 1998. Formulation of a three dimensional shelf edge model and its application to internal tide generation. Continental Shelf Research 18, 405–440.

ELSEVIER

Deep-Sea Research II 51 (2004) 3009–3028

DEEP-SEA RESEARCH
PART II

www.elsevier.com/locate/dsr2

Spatial structure of anticyclonic eddies in the Algerian basin (Mediterranean Sea) analyzed using the Okubo–Weiss parameter

Jordi Isern-Fontanet[a,*], Jordi Font[a], Emilio García-Ladona[a], Mikhail Emelianov[a], Claude Millot[b], Isabelle Taupier-Letage[b]

[a]*Institut de Ciències del Mar, CMIMA-CSIC, Passeig Marítim de la Barceloneta 37-49, 08003 Barcelona, Spain*
[b]*Laboratoire d'Océanographie et de Biogéochimie, COM-CNRS, BP 330, 83507 La-Seyne-sur-Mer, France*

Abstract

Three Algerian eddies (open sea big anticyclonic eddies in the Algerian basin, western Mediterranean sea) are studied in detail using altimetric and in situ data. To characterize their spatial structure we make use of the Okubo–Weiss parameter, which allows to separate the flow into vorticity-dominated and deformation-dominated regions. The application of this parameter to geostrophic velocity fields obtained from Sea Level Anomaly maps show that the observed eddies have a spatial structure in close resemblance to that found in coherent vortices of two-dimensional turbulence. These eddies appear to have a rather coherent vorticity-dominated core region surrounded by an outer region dominated by deformation with a very complicated and variable in time structure. This analysis is also applied to CTD data collected across three eddies sampled between 1997 and 1998. From these data eddy cores are equivalently defined as the interior region enclosed by the maximum of azimuthal velocity, and are consistent with the horizontal structures seen in SLA maps. The distinction between the core and the outer region (circulation cell) is particularly important to understand the role of transport and stirring of properties associated to eddies, as illustrated by the observed salinity structure, some selected SST images, and drifter trajectories. The core is characterized by very small isopycnal salinity gradients, while the surrounding cell is characterized by a strong slope of isohalines with respect to isopycnals. Persistent stagnation points are observed around the eddy and correspond to regions of high values of deformation with respect to vorticity. These regions allow analyzing characteristic patterns seen in the corresponding SST images. In addition, drifters trajectories exhibit very different behavior when crossing such regions.
© 2004 Elsevier Ltd. All rights reserved.

1. Introduction

Several studies have investigated the redistribution of properties in the ocean associated with intense mesoscale activity (e.g. Wilkin and

*Corresponding author.
E-mail address: jisern@icm.csic.es (J. Isern-Fontanet).

0967-0645/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dsr2.2004.09.013

Morrow, 1994; Abraham and Bowen, 2002). This intense eddy activity increases mixing in the surface layer and is mainly due to the presence of energetic and long lived structures generically named eddies (Robinson, 1983; Kamenkovich et al., 1986). In the Algerian Basin, western Mediterranean Sea, large anticyclonic open sea eddies (see Fig. 1A) have been repeatedly observed (Millot, 1985; Taupier-Letage and Millot, 1988; Benzohra and Millot, 1995b, Obaton et al., 2000; Salas et al., 2002; Ruiz et al., 2002). These mesoscale structures, named Algerian Eddies (AEs), are generated, a few times per year, as the result of an instability of the along slope Algerian current. They have variable diameters of about 100–200 km, vertical extents from 100s to 1000s of m, and can have lifetimes of several months, up to near 3 years (Millot et al., 1997; Puillat et al., 2002). The common formation and presence of these big eddies has a large impact on the surface and intermediate circulation in the western Mediterranean (Millot, 1987; Send et al., 1999; Fuda et al., 2000). They can interact with other AEs in the mid-basin, or with their parent current, sometimes leading it to be diverted seaward from the coast (Taupier-Letage and Millot, 1988; Benzohra and Millot, 1995a).

The role of these eddies on the redistribution of properties in the Algerian basin is related with the northward spreading of water mass properties that should initially be transported eastwards by the Algerian current. Evidences of a meridional momentum transport have been observed through velocities obtained from satellite Sea Level Anomaly (SLA) maps (Iudicone et al., 1998), mean climatological maps of the surface circulation (Ovchinikov, 1966; Tziperman and Malanotte-Rizzoli, 1991) and statistical analysis of drifters (Salas et al., 2001). The signature is a meridionally enlarged and smoothed current pattern towards the central part of the Algerian basin. Similar features are observed in all climatological atlas of water properties (Picco, 1990; Brasseur et al., 1996) and in numerical simulations of the mean circulation (Herbaut et al., 1997). As an example, in Fig. 1B a mean Sea Surface Temperature (SST) image corresponding to the first semester of 1998 shows the relatively smoothed meridional gradient of temperature to the north. This contrasts with the SST image of the Algerian basin (Fig. 1A) with complex patterns associated to three big anticyclons occupying most of the basin. The effective mixing is mainly due to the non-local transport of these eddies as they move through the whole basin (see Fig. 2 and Olson, 1991; Bouzinac et al., 1998; Puillat et al., 2002).

While theoretical and numerical simulations have shown the importance of eddies in the local dispersion of particles and their role in the evolution of tracer concentrations (e.g. Provenzale, 1999), less is known on the mixing and the stirring properties associated to these individual structures in the ocean. Dewar and Flierl (1985) analyzed with some detail the link between the stream function topology and the dispersion of a passive tracer for simple eddy models, and compared with them some features seen in SST images. In general, eddies have different dynamical properties between the inner and outer parts of their structure. While the interior part of the eddy, or core, transports tracers within it, the outer produces intense stirring of the ambient environment (Olson, 1991). Analysis of numerical results on two-dimensional turbulent flows have proposed the use of a general criterion based on the Okubo–Weiss parameter (Okubo, 1970; Weiss, 1991) to determine the detailed spatial structure of an eddy and to identify the regions with different mixing properties (McWilliams, 1984; Elhmaïdi et al., 1993).

In a previous paper (Isern-Fontanet et al., 2003) some of us have proposed to use a criterion to detect eddies from SLA maps based on the partition of the flow using the Okubo–Weiss parameter. In the present paper, we show, combining several data sets, the contribution of this criterion to understand the associated stirring properties. Data are described in Section 2 and consist on SLA maps, SST images, hydrographic measurements, and trajectories of drifters buoys. The definition of the Okubo–Weiss parameter and its application to these data are explained in detail in Section 3. The relation between stirring and mixing properties of the regions partitioned given by the parameter is analyzed in Section 4. We discuss the results and limitations of the approach

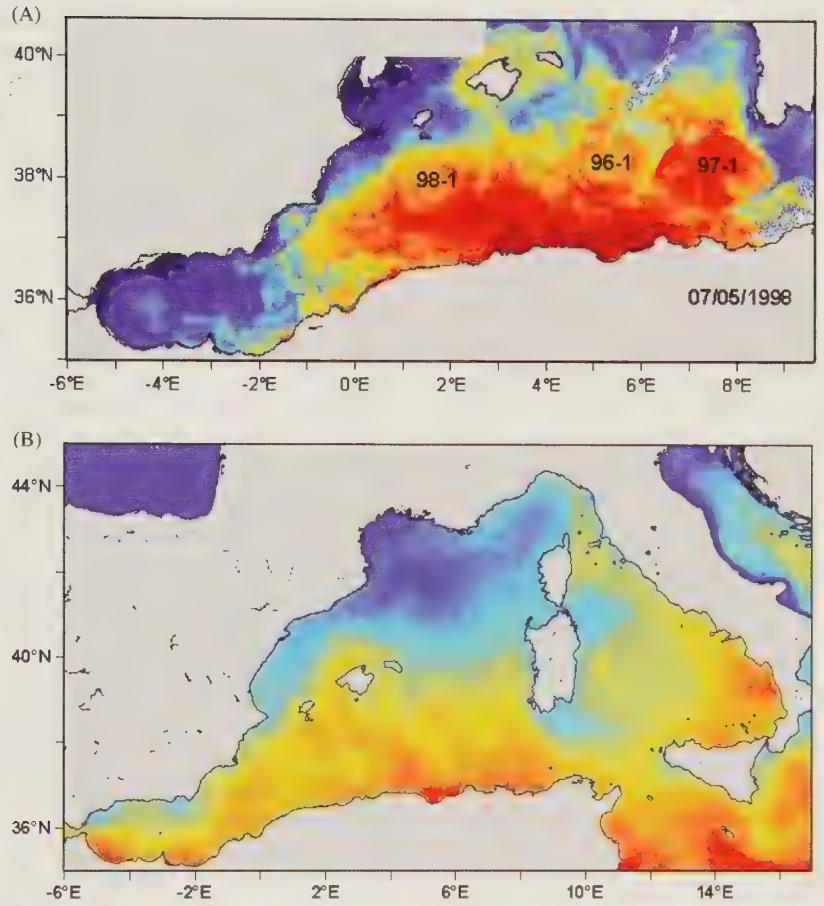


Fig. 1. (A) SST image for May 7, 1998 with the location of 96-1, 97-1 and 98-1 eddies. (B) Mean Sea Surface Temperature for the first semester of 1998. Temperature increases from blue to red, the contrast is adjusted specifically for the image to enhance the signatures of the eddies.

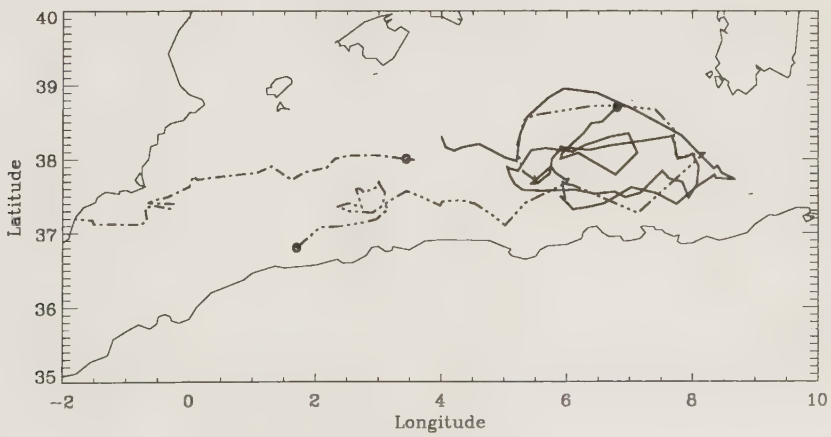


Fig. 2. Observed trajectories of 96-1 (solid line), 97-1 (dot-dot-dot-dash line) and 98-1 (dot-solid) eddies. Circles indicate the initial position observed in SLA maps. These trajectories were obtained using the methodology described in Isern-Fontanet et al. (2003).

in Section 5 and finally we summarize the main results.

2. Data sets

2.1. Satellite data

In the Mediterranean Sea, despite the weak signal intensity and the coarse space and time resolution of the altimetric tracks, several studies have shown the reliability of the altimetric data to analyze its dynamics and particularly in the Algerian basin (Vazquez-Cuervo et al., 1996; Vignudelli, 1997; Ayoub et al., 1998; Iudicone et al., 1998; Bouzinac et al., 1998; Larnicol et al., 2002; Font et al., 2004). Here we use SLA maps regularly produced by Collecte Localisation Satellites (CLS) in Toulouse, France (http://www.cls.fr) from a combination of both *ERS* and *TOPEX/Poseidon* altimeters measurements. These maps are processed including usual corrections (tides, inverse barometer, etc.), a Sea State empirical model with four parameters derived from a global crossover analysis (Gaspar et al., 1994) and with improved *ERS* orbits using *TOPEX/Poseidon* as a reference (AVISO User Handbook, 1996; Le Traon and Ogor, 1998). Sea level anomalies are used instead of absolute sea surface height due to the still unprecise determination of the marine geoid. SLA are regularly produced by subtracting a 4-year mean value from the actual measurements and, prior to the analysis, data are low-pass filtered using a 35 km median filter and a Lanczos filter with a cutoff wavelength of 42 km in order to reduce altimetric noise (Larnicol et al., 1995). SLA maps are finally built on a regular grid of 0.2×0.2 degrees every 10 days using an improved space/time objective analysis method, which takes into account long wavelength errors (Le Traon et al., 1998).

The infrared satellite images used here come from the NOAA/AVHRR sensor, have been processed at the German Remote Sensing Data Center, and are freely available through its database server (http://www.dfd.dlr.de). Images are either processed up to SST and combined temporally to daily, weekly and monthly composites, or are plain AVHRR channel 4 of AVHRR corresponding to relative brightness temperature (see details on the processing in the above web site). The images here used are selected pictures to be almost simultaneous with the SLA maps. They are part of the series of images used to track eddies in the Algerian basin by Puillat et al. (2002).

2.2. In situ data

In connection with the EU MATER (*MAss Transfer and Ecosystem Response*) project, two field programs were carried out in the Algerian basin to study the dynamics of AEs. ELISA (*Eddies and Leddies Interdisciplinary Study in the Algerian basin*) in the eastern part (Taupier-Letage et al., 2003) and ALGERS (*the use of ERS sensors for the study of the dynamics of Atlantic water in the ALGerian basin*) in the western part (Font et al., 1998). Three big anticyclonic eddies, labeled 96-1, 97-1 and 98-1 (Puillat et al., 2002), were detected and sampled with CTD casts between 1997 and 1998 during ELISA1, ALGERS98 and ELISA4 cruises, respectively. These three eddies were particularly long lived and coherent structures with a clear signature in infrared images (Puillat et al., 2002; Salas, 2003) and altimetry (Isern-Fontanet et al., 2003; Font et al., 2004).

Casts during ELISA1 and ELISA4 campaigns were performed from the French RV Le Suroit with a SBE 911 plus CTD profiler from Sea-Bird Electronics, Inc. During the ALGERS98 campaign the CTD data were obtained from the Spanish RV Hespérides with a Neil Brown Mark-III probe. All data were preprocessed with the SEASOFT (v. 4.233, Sea-Bird Electronics, Inc.) software package. Final profiles were interpolated with a vertical resolution of 1 dbar. More details on data processing can be found in Ruiz et al. (2002) and Taupier-Letage et al. (2003). From all the available CTD profiles we have selected those that crossed an eddy close to one of its diameters (Fig. 3). While during ELISA4 and ALGERS98 the sections crossed completely the eddy, in ELISA1 the CTD section cut the eddy through a sector at the north and not passing close to the eddy center (Taupier-Letage et al., 2003). Cross-sections of geostrophic velocities for all the three transects were derived from

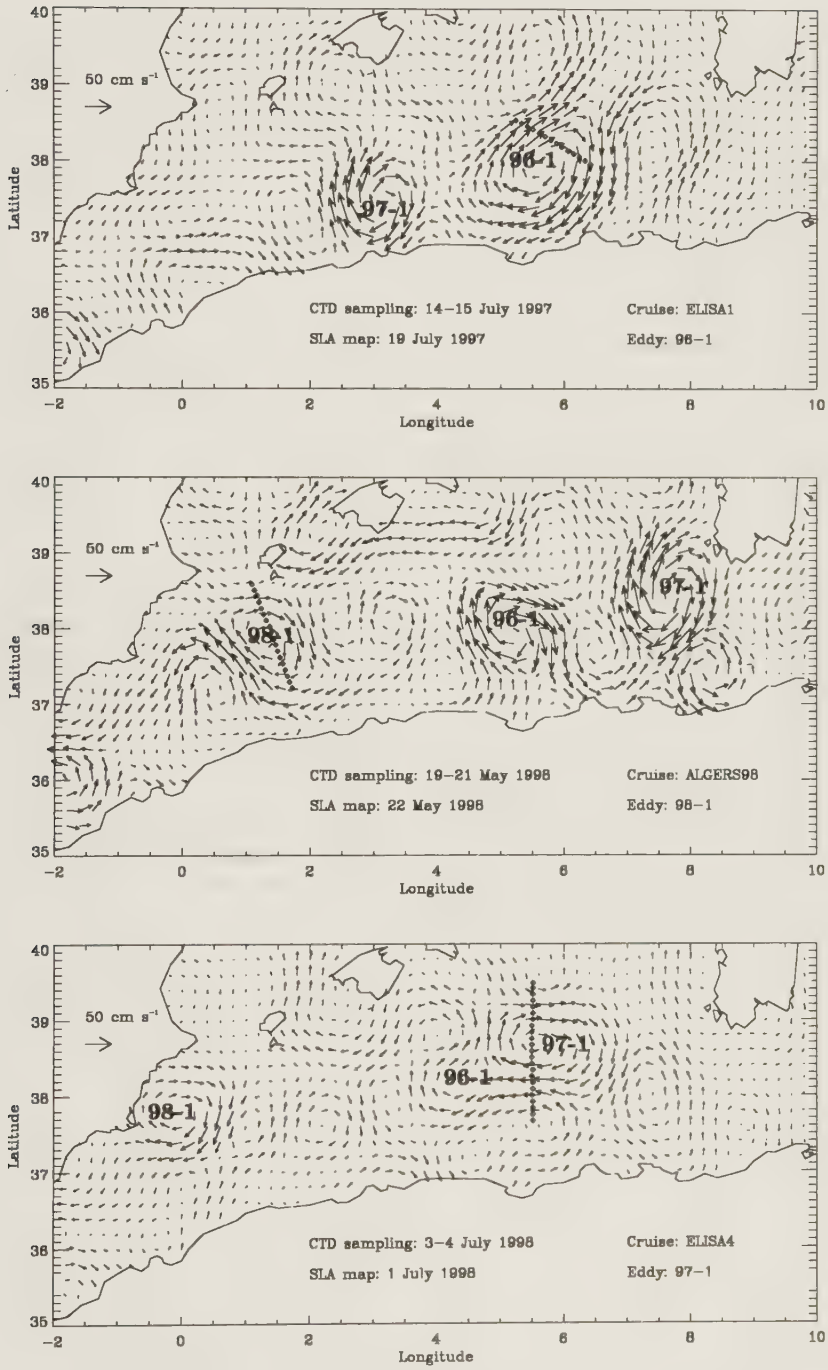


Fig. 3. Geostrophic velocities derived from SLA maps close to CTD sampling (diamonds) of eddies 96-1, 97-1 and 98-1.

dynamic heights with a reference level at 1000 dbar, similar to previous works in the region (Benzohra and Millot, 1995b). Resulting geostrophic velocities showed that the Rossby numbers for these eddies were $Ro = 0.13$ for 96-1, $Ro = 0.11$ for 97-1 and $Ro = 0.09$ for 98-1. Then, the gradient flow approach was used instead of direct geostrophic velocities following Olson (1991).

In ALGERS98 cruise, 15 surface drifting buoys (drogued with a 10 m long holy sock at a mean depth of 10 m) were released across the eddy at the positions of the CTD casts (Fig. 3). The drifters were tracked during a variable number of days, from 12 to 148, until transmission stopped due to failure of the system at open sea or after having gone ashore (Salas, 2003; Font et al., 2004).

3. The Okubo–Weiss parameter

3.1. Definition and properties

Several studies in the past have revealed the close link between the local topology of the flow field, the stirring and mixing processes, and the dispersion of Lagrangian particles in complex flows. Okubo in 1970 analyzed the dispersion in the ocean surface by looking at the shape of the trajectories of two particles initially located close to each other. By means of a linear analysis he found that the shape of the possible particle trajectories is characterized by the eigenvalues of the velocity gradient tensor. Later Weiss in 1981 arrived to an equivalent criterion analyzing the enstrophy dynamics in two-dimensional turbulent flows. In both approaches, the topology of particle trajectories and the behavior time evolution vorticity gradients is determined by the sign of

$$W = s_n^2 + s_s^2 - \omega^2, \tag{1}$$

where s_n, s_s and ω are the normal and the shear components of strain and the relative vorticity of the flow defined respectively by

$$s_n = \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}, \quad s_s = \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}, \quad \omega = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}.$$

The parameter W, named after them as Okubo–Weiss parameter, allows to separate the flow into different regions: a vorticity-dominated region ($W < -W_0$), a strain-dominated region ($W > W_0$) and a background field, characterized by small positive and negative values of W ($|W| \leqslant W_0$), where $W_0 = 0.2\sigma_W$ with σ_W being the standard deviation in the whole domain, in our case the Mediterranean Sea (Bracco et al., 2000; Pasquero et al., 2001). When applying this criterion of flow partition it can be observed that the general structure of an eddy consists in a vorticity-dominated inner region, the eddy core, surrounded by a strain-dominated region, the circulation cell (Elhmaïdi et al., 1993). The core edge can then be identified as the closed lines with $W = 0$. This separation of the field in terms of the sign of W has been proven to be a robust criterion for educing eddy cores from complex fluid flows (Jeong and Hussain, 1995; Pasquero et al., 2001; Isern-Fontanet et al., 2003). Furthermore, the boundary of the core constitutes a barrier to the exchange of particles with the surrounding cell, then particles trapped inside the eddy core remain there unless the eddy is destroyed or they are ejected through a filamentation process. In the circulation cell, on the other hand, there exist an intense stirring and exchange of particles with the background field. During the dispersion process, neutral particles exhibit a tendency to concentrate in such region (Provenzale, 1999).

For two-dimensional flow fields, as those obtained from SLA maps, the criterion is straightforward implemented by computing the velocity field and its derivatives assuming geostrophic balance

$$u = -\frac{g}{f}\frac{\partial h'}{\partial y}, \quad v = \frac{g}{f}\frac{\partial h'}{\partial x},$$

where h' is the Sea Level Anomaly, g gravity, and f the Coriolis parameter. However, for velocity transects, as those obtained from aligned CTD measurements, the criterion is not so easily implemented since it requires some assumptions on the shape of the eddy. In polar coordinates the Okubo–Weiss parameter, assuming a horizontally non-divergent flow, is given by

$$W = 4\left[\left(\frac{\partial v_r}{\partial r}\right)^2 + \frac{1}{r}\left(\frac{\partial v_r}{\partial \theta}\right)\left(\frac{\partial v_\theta}{\partial r}\right) - \frac{1}{r}\left(\frac{\partial v_\theta^2}{\partial r}\right)\right], \tag{2}$$

where v_r and v_θ are the radial and azimuthal velocities, r is the radius and θ is the azimuthal angle. A common assumption is that the eddy has axial symmetry. Then the velocity field is given by

$$v_r = 0, \quad v_\theta = V(r)$$

and then the Okubo–Weiss parameter is

$$W = -4\frac{1}{r}\left(\frac{\partial V^2}{\partial r}\right). \tag{3}$$

This implies that the boundary of the core ($W = 0$) corresponds to the radial location of the maximum of velocities. In a general case, the line of maximum kinetic energy is close to the line of $W = 0$ but they do not necessarily coincide, as is seen from Eq. (2).

3.2. Satellite observations

As illustration, in Fig. 4A we have represented the horizontal distribution of W derived from a SLA map corresponding to July 7, 1996. A color scale has been selected so that blue colors represent eddy cores ($W < -0.2\sigma_W$), green represents the background field ($|W| \leqslant 0.2\sigma_W$) and yellow and red colors the deformation-dominated regions ($W > 0.2\sigma_W$). When this figure is compared with a SST image (Fig. 4D) it can be observed that the core region of strong negative values of W is consistent with surface patterns labeled as 96-1, as well as the position of other eddies. One exception is the eddy located south of Mallorca island (38.5°N, 3°E) in SST image, which is not seen in the W field. However, SLA maps contain more eddy cores than those identified by visual inspection in infrared images, although it is not possible to confirm whether or not are real eddies (see Isern-Fontanet et al. (2004) for a detailed discussion). The structure of 96-1 in terms of W (Fig. 4A) displays a nearly axisymmetrical core surrounded by a region of intense deformation and irregular shape. A zonal cross-section along one of its diameters (Fig. 4B and C) shows that W (red line) takes strong negative values in the inner part (the eddy core) and becomes positive (the circulation cell) after crossing the radial position of the kinetic energy (blue line) maximum. This structure is very similar to the structure of two-dimensional vortices as those observed in turbulence (for instance compare it with Fig. 3 of Elhmaïdi et al., 1993). To provide an evidence of the evolution of these eddies from the point of view of W, Figs. 5A–C show the spatial distribution of the parameter for three consecutive SLA maps close to ELISA1 cruise. It can be appreciated that while eddy cores are rather coherent structures in both space and time the circulation cells are much more irregular in space and variable in time.

The diameter of eddy cores can be estimated from SLA by calculating the diameter of an equivalent circular core with the same area. Observations show that 96-1 has a diameter that ranges from 95 km on 9 July 1997 to 116 km on 19 July 1997, 97-1 ranges from 116 km on 1 July to 96 km on 11 July 1998 and 98-1 from 97 to 77 km. Notice, however, that the core of 97-1 was very elongated on 1 July 1998 as observed from Fig. 3. For 96-1 the northern boundary is located between CTD 31 and 32, and the southern limit is beyond the CTD casts. The limits for 97-1 are close to CTD 235 in the north and between CTD 244 and 245 in the south. For 98-1 the limits are close to CTD 5 in the north and between CTD 14 and 15 in the south.

From a sequence of maps as the one shown in Fig. 5 the trajectories of these eddies can be recovered (as explained in detail by Isern-Fontanet et al., 2003). Once identified in one map they are tracked by looking at nearby regions of strongly negative values of W in the next map. Using this method, 96-1 was first identified in SLA maps on 19/2/1996 and lived until 21/7/1998. Eddy 97-1 from 20/5/1997 until 18/9/1998 and 98-1 from 22/2/1998 until 7/11/1998. These results give lifetimes of 883, 486 and 139 days, respectively. In situ measurements were made when 96-1 was 512 *days* old (ELISA1), 97-1 was 410 days old (ELISA4) and 98-1 was 88 *days* old (ALGERS98).

3.3. In situ observations

The consistency between the spatial distinction between eddy regions in SLA maps as determined by W, and the characteristics of the vertical structure of these eddies, is checked by representing vertical sections of σ_t and gradient velocity from the CTD measurements that approximately

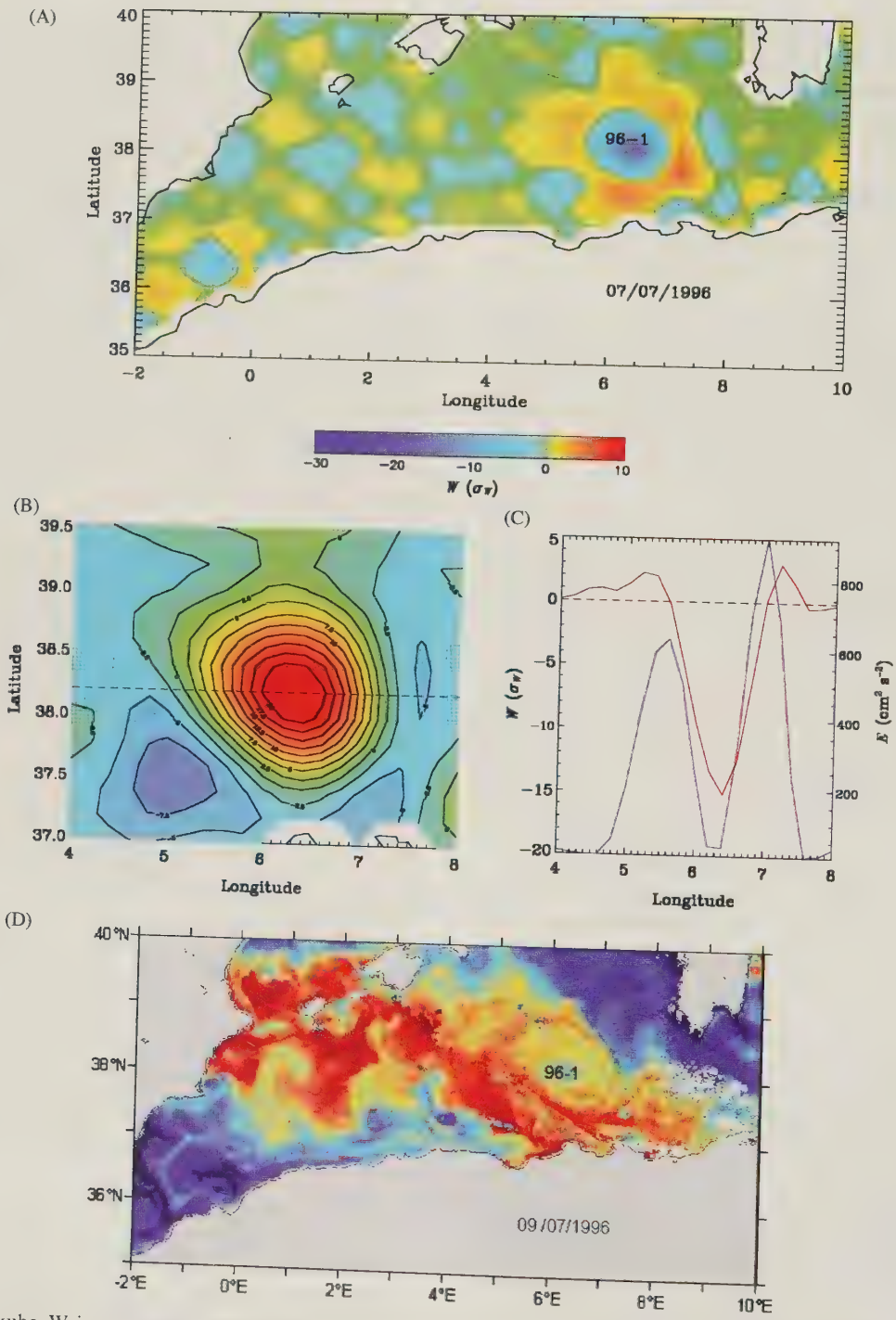


Fig. 4. The Okubo–Weiss parameter computed for the 96-1 eddy from SLA maps. (A) General view of the distribution of the Okubo–Weiss parameter on 7/7/1996. The color scale has been chosen to enhance the values above and below the line $W = 0.2\sigma_W$ ($\sigma_W = 2.44 \times 10^{-11}/s^2$). (B) Sea level anomalies (cm) corresponding to the 96-1 eddy with the transect drawn in figure C (dashed line). (C) Longitudinal profile of the Okubo–Weiss parameter (red line) and kinetic energy (blue line) across the eddy. (D) SST image corresponding to July 9, 1996. Temperature increases from blue to red, the contrast is adjusted specifically for the image to enhance the signatures of the eddies.

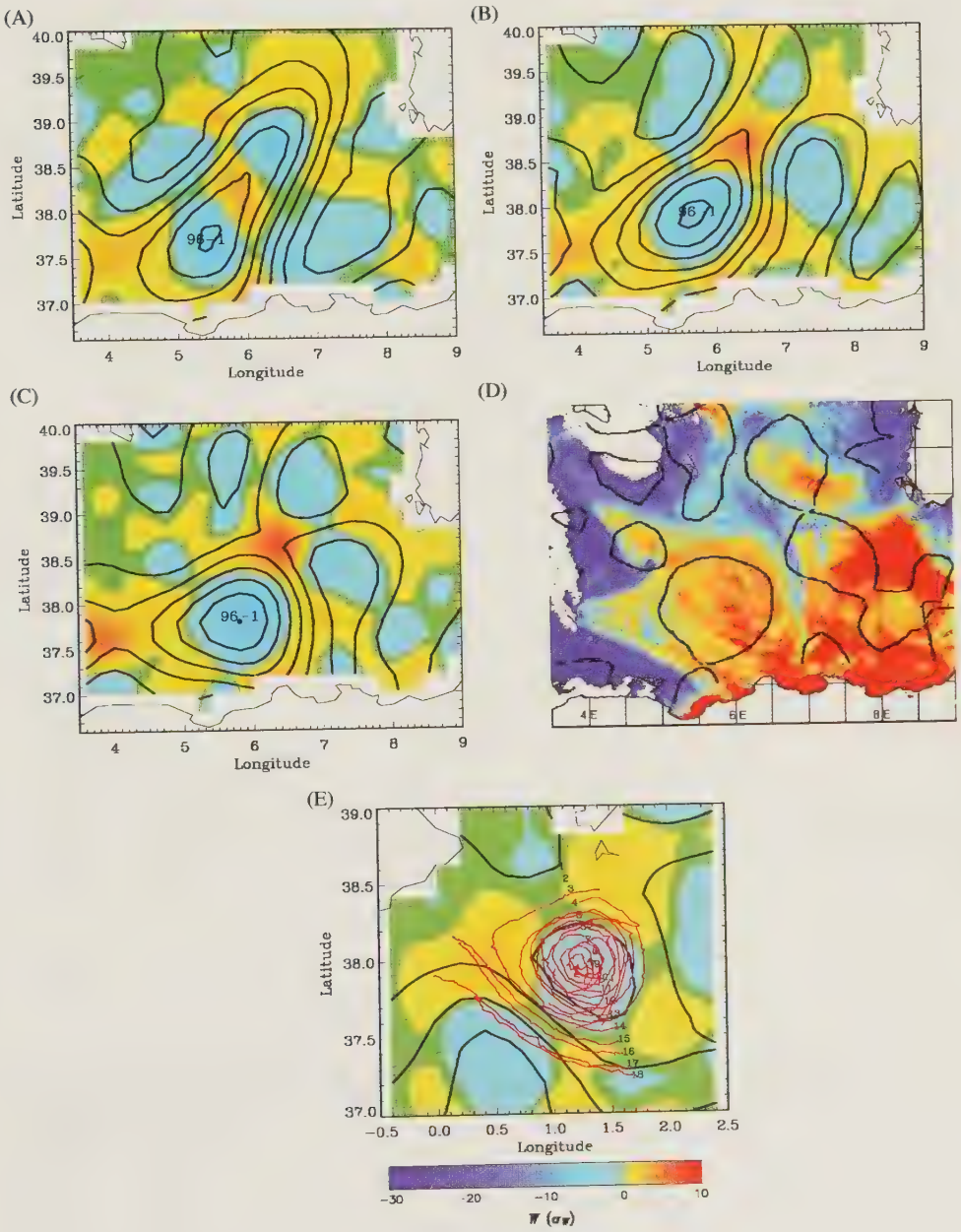


Fig. 5. Distribution of the Okubo–Weiss parameter and streaklines (black) derived from SLA maps for 96-1 close to ELISA1 cruise: (A) July 9, 1997 ($\sigma_W = 2.64 \times 10^{-11}/s^2$), (B) July 19, 1997 ($\sigma_W = 3.39 \times 10^{-11}/s^2$), (C) July 29, 1997 ($\sigma_W = 3.31 \times 10^{-11}/s^2$). (D) SST image corresponding to July 30, 1997 with lines of $W = 0$ (black). (E) Distribution of the Okubo–Weiss parameter and streak lines derived from SLA maps for 98-1 close to ALGERS98 cruise, May 22, 1998 ($\sigma_W = 3.42 \times 10^{-11}/s^2$) and initial trajectories of drifters (red lines) launched during the cruise. Numbers indicates CTD casts. Temperature increases from blue to red, the contrast is adjusted specifically for the image to enhance the signatures of the eddies.

crossed 96-1, 97-1 and 98-1 cores (Figs. 6 and 7). The vertical structures are characteristic of anticyclonic eddies with the center located between casts 27–29, 237–238 and 8–10 for 96-1, 97-1 and 98-1, respectively. As expected the vertical structures are similar for both temperature and salinity (see Ruiz et al., 2002; Taupier-Letage et al., 2003) except for the presence of a shallow thermocline located within the upper 50 m, clearly reflected in the σ_t sections of 96-1 and 97-1, and characteristic of the Mediterranean summer season. The vertical extent of the baroclinic structure

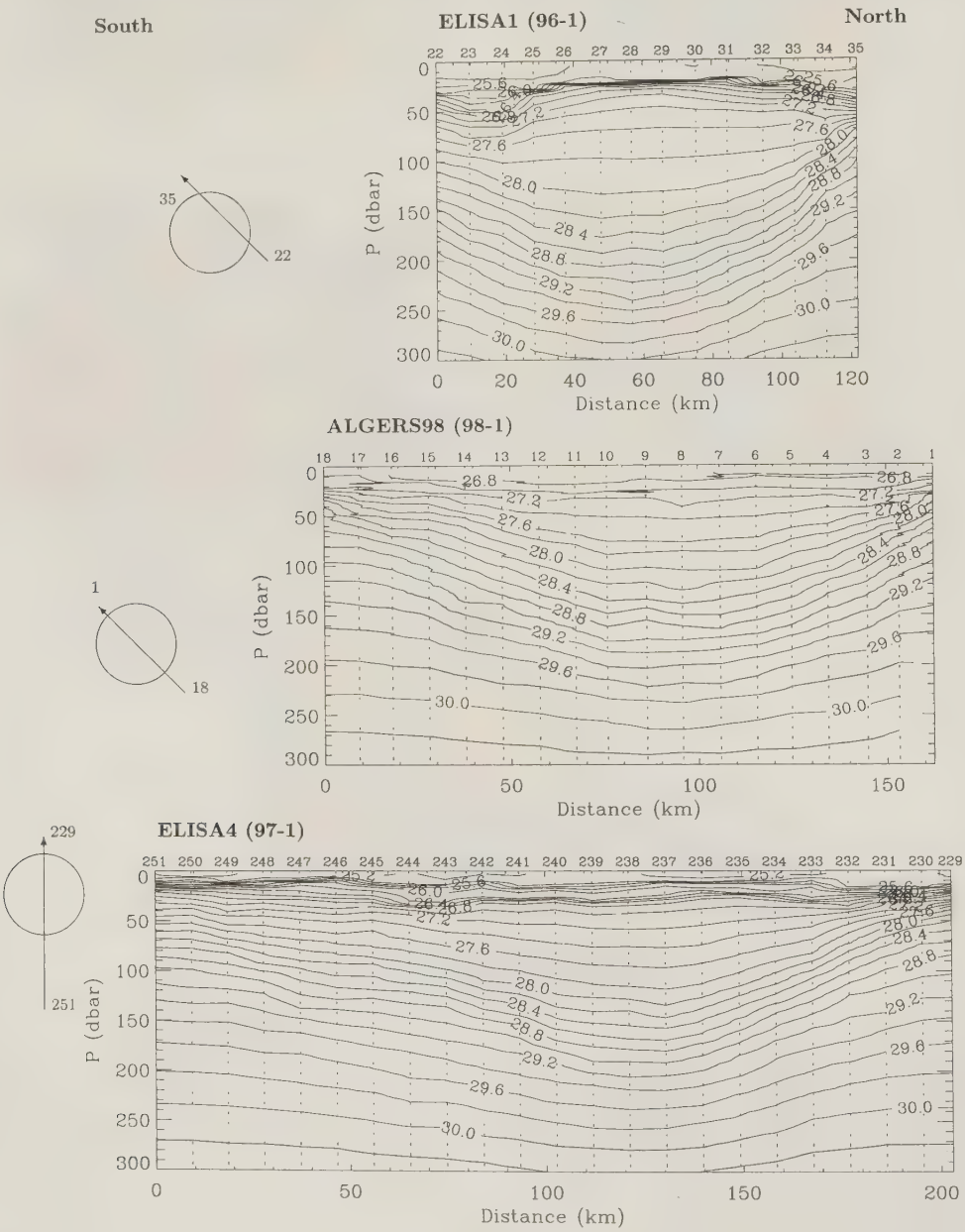


Fig. 6. Transect of σ_t (kg/m^3) corresponding to the CTD casts shown in Fig. 3. Dotted lines are drawn at the CTD location. Numbers identify CTD casts.

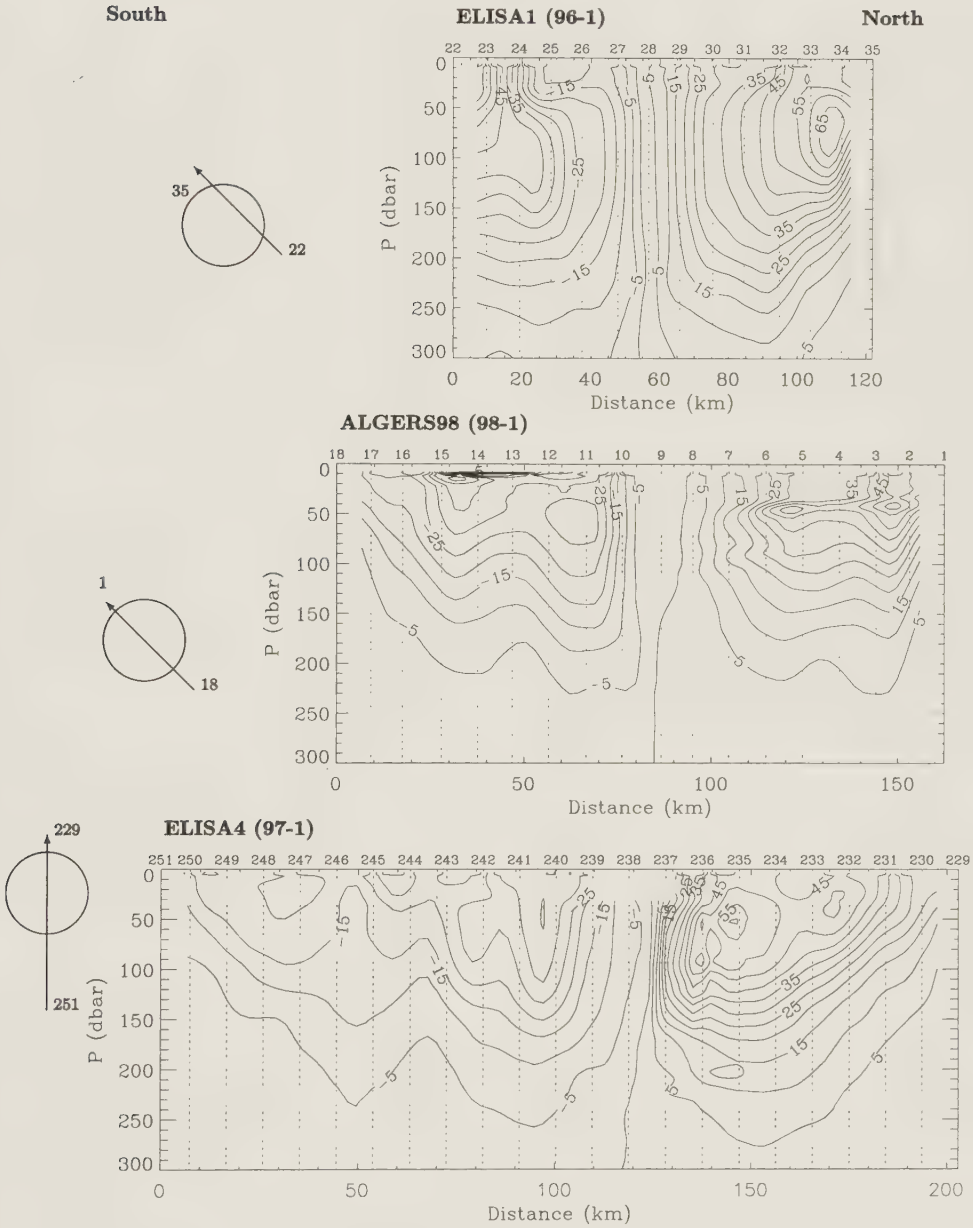


Fig. 7. Transect of the gradient flow velocities (cm/s) corresponding to Fig. 3.

is about 250–300 m and very similar for the three eddies. Gradient velocities around 50–60 cm/s can be found in the upper part, while velocities are lower than 5 cm/s below 300 m depth. For 96-1 the spatial distribution is rather symmetric with a well defined velocity extreme in the north (between casts 33 and 34, Fig. 7) around 65 cm/s. In the south the velocity extreme was not sampled although a relative maximum can be observed around cast 24. As pointed before, some experimental evidences confirm that this CTD section was shifted to the north from the center of the

eddy structure (Taupier-Letage et al., 2003). Instead, the sections through 97-1 and 98-1 seem to capture the full horizontal extension of the eddies. Furthermore, CTD sampling in the southern part of both eddies is even clearly beyond their spatial extent. However, the spatial structure of these velocity sections has a certain complexity.

Eq. (3) has been applied to the gradient velocities shown in Fig. 7 and revealed that the resulting distribution of W is more complex than the one observed in SLA maps. However, since results were noisy, we restricted our analysis to the identification of core boundaries, and left a detailed analysis of the distribution of W in these eddies for a future study. As seen before, the identification of core boundaries, under the assumption of axial symmetry, reduces to the identification of velocity extrema. The application of this criterion to 96-1 (Fig. 7) shows that the northern boundary of the core changes with depth. Above 50 dbar it is located between CTD casts 32 and 33, while below it is located between 33 and 34. Below 130 dbar the core reduces its size and the boundary is located between casts 31 and 32. At the southern limit the boundary is located beyond the measurements. However, at a depth of about 100 dbar there is a relative maximum of velocities between cast 24 and 25. These results suggest that the diameter of the eddy core ranges from 104 to 65 km. The same analysis is applied to 98-1. It can be observed that the northern edge has two velocity extrema between CTD 2 and 3 and 5 and 4 for depths of 40 dbar. At the southern part of the eddy two velocity extrema can be found, one between casts 11 and 12 at a depth of 50 dbar and the other between 14 and 15 at the surface. Taking the inner maxima the eddy diameter is about 60 km while with the outer it is about 115 km. Finally, the 97-1 velocity structure is more complicated. The presence of several relative velocity maxima can be identified. In the northern part of the eddy there are two relative maxima in the upper layers, one between CTD 236 and 235 and the other between 232 and 233. In the southern limit velocity maximum is located between casts 240 and 241. This diameter ranges from 50 to 75 km.

4. Relation with mixing

As previously outlined, the distribution of W allows identifying regions with different mixing properties. In particular, for an eddy the core is characterized by low radial mixing, and the circulation cell by intense stirring with surrounding waters. We analyze this here through the distribution of salinity, obtained from CTD casts, and SST images. The information provided by these variables is complemented by the drifting buoys trajectories observed during ALGERS'98.

4.1. SST distribution and drifters

As shown by Dewar and Flierl (1985), in a frame moving at a constant velocity the region of particle trapping is the region enclosed by the critical streakline. Streaklines associated with a stream function ψ, representing an eddy propagating at a constant velocity (c_x, c_y), are given by

$$\chi = \psi - c_x y + c_y x. \tag{4}$$

Dewar and Flierl (1985) applied it to simple models of eddies and to numerical simulations of a propagating eddy. They found that the trapping region is beyond the radial location of the maximum of velocities, which approximately defines the boundary of the eddy core in quasi axisymmetric flows.

The estimation of χ is straightforward from altimetry if it is assumed that the eddy is steadily propagating. We have estimated the velocity of the eddy from two sequential SLA maps and computed the streaklines (Figs. 5A–C). The critical streakline should contain a stagnation point located between the closed contours around the eddy core and the first non-closed open contour. For example, in Figs. 5A–C it can be appreciated an almost stationary region observed around 37.5°N, 4°E, and in Figs. 5B, C also around 38.75°N and 6.5°E. Both regions have a strong hyperbolic character because they are dominated by deformation, that is positive values of W. SST image (Fig. 5D) exhibits two interesting features. On one hand, a pattern of relatively high-temperature clearly deformed to the west with a characteristic shape similar to the critical

streakline. The end of this lobular structure coincides with the region of great deformation. On the other hand, over the area of the second saddle point at the north of the eddy, a filament-like structure of relatively cold waters penetrates to the south surrounded by warmer waters and around the eddy core region.

During ALGERS98 several surface drifters were deployed at the same positions of CTD casts (Font et al., 2004). Their trajectories are overploted here on the W map together with streaklines (Fig. 5E). The drifters deployed between CTD 5 and 14, which correspond to the eddy core, completed at least one loop around the eddy, as well as those drifters that where deployed at the positions of CTD 4, 15 and 16. Drifters deployed at CTD 17 and 18 escaped and did not complete any loop. In the figure the critical streakline is approximately situated between casts 14 and 16. In these locations two buoys started drifting directly to the region of the saddle point characterized by high values of deformation, and then moved eastwards until completing a loop (not shown in the figure). The drifter released in 14 also did a similar initial trajectory but after passing the saddle point region continued to loop around the eddy (see Font et al., 2004 for detailed trajectories).

4.2. Salinity distribution

The implications for the mixing properties associated to the core and circulation cell for such eddies can be analyzed by considering the salinity in the upper layer as a tracer. Because the motion tends to be constrained to isopycnal surfaces, the spatial partition between core and circulation cell should be more adequate in terms of isopycnic coordinates. In order to analyze the vertical cross-sections the data were interpolated to isopycnal surfaces with a separation of $\Delta\rho = 0.05\,\mathrm{kg/m^3}$. A common characteristic of the three eddies is the presence of a center with salinity about 37 surrounded by saltier waters with salinity of the order of 38. The T-S diagrams of CTD casts in the center of these eddies (Fig. 8) show the isohaline nature of the eddy core down to the 27.0–27.5 $\mathrm{kg/m^3}$ isopycnal depending on the eddy. The contrast between the less saline core with respect to the surrounding waters is particularly enhanced in this surface layer. Below this isopycnal the salinity differences with respect to surrounding waters are smaller, making difficult to clearly distinguish between the core and the circulation cell.

Fig. 9 shows the isopycnal distribution of salinity (thick lines) corresponding to transects shown in Figs. 6 and 7. Gradient velocity (thin lines) have been overploted in order to make easier its comparison with salinity. For 96-1, salinity distribution inside the core is homogeneous with some well radially localized low-salinity anomalies for the upper part of the eddy, above $28\,\mathrm{kg/m^3}$. These salinity anomalies, located around $27.5\,\mathrm{kg/m^3}$, are observed at both sides of the eddy (casts 31–34 and 23–24) in the inner part of the core boundary. They can also be observed in another transect (not shown) between casts 40 and 45, which suggests that they could have a toroidal structure. Unfortunately, we do not have additional data to confirm this point. On the other hand, beyond the northern velocity maximum (casts 33 and 34) there are intense gradients of salinity. In the south, there is a small salinity gradient located between casts 24 and 25 around the $28\,\mathrm{kg/m^3}$ isopycnal. This isopycnal variation of salinity is approximately located at the position of the maximum of velocities reported above. For eddy 98-1 the low salinity layer in the core is thiner and the analysis is more difficult. However, the salinity distribution shows the presence of strong gradients to the north of the northern velocity maximum (casts 3 and 2). Between CTD 4 and 6 it seems to be some salinity gradient that would correspond to the location of the inner maximum. At the southern part of the transect the gradients seem to be stronger at the south of the extreme located between casts 12 and 11. In 97-1 core the isohalines are predominantly flat, although there is some curvature that increases with increasing radius and becomes very steep after crossing the outer velocity maxima. In such a complex geometry the situation is not clear.

5. Discussion

We have evidenced the close link between the local spatial structure of the marine eddies, its

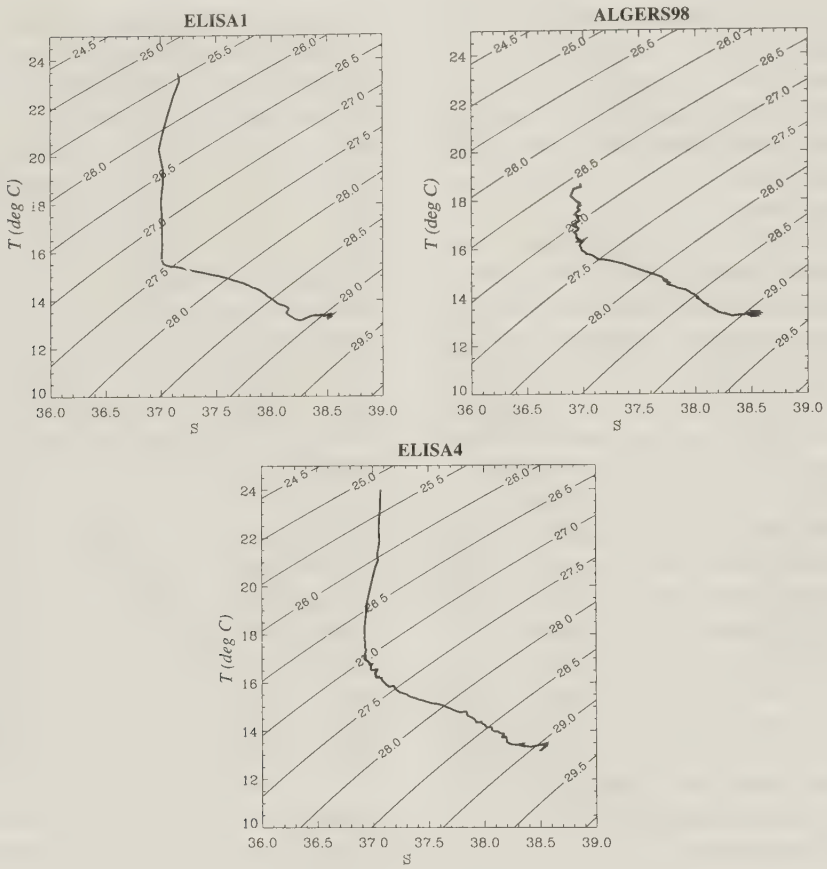


Fig. 8. T-S diagrams of the centers of eddies 96-1 (cast 27), 98-1 (cast 10) and 97-1 (cast 240). Lines of constant σ_t (sloping lines) are given in kg/m^3.

associated stirring properties and the spatial distribution of tracers exemplified by temperature and salinity. Unfortunately, the study is somehow restricted to find precise quantitative results essentially due to the limitations associated with the data used. Several points must be considered and discussed to evaluate the validity of the assumptions.

The basic approach to analyze the spatial structure of these eddies has been the use of W as a way to separate the different topological regions of the flow. The close resemblance of the profiles of W when an eddy is detected, and the pattern found for a vortex in two-dimensional turbulence simulations is quite remarkable (Fig. 3 in Elhmaïdi et al., 1993). This agreement between both flow fields is because, that for ocean flows at reasonably large scales, motion is almost divergenceless and tends to be constrained to isopycnal surfaces that are near horizontal. This makes it reasonable to consider a similar spatial partition applied to the flow along isopycnal surfaces. However, the flow partition in terms of straining and vorticity dominated regions, as given by the values of W, has some limitations. Two-dimensional decaying turbulence studies were proved that the validity of W is restricted to regions within eddy cores and near stagnation points (Basdevant and Philipovitch, 1994), and the partition may be rigorously addressed in

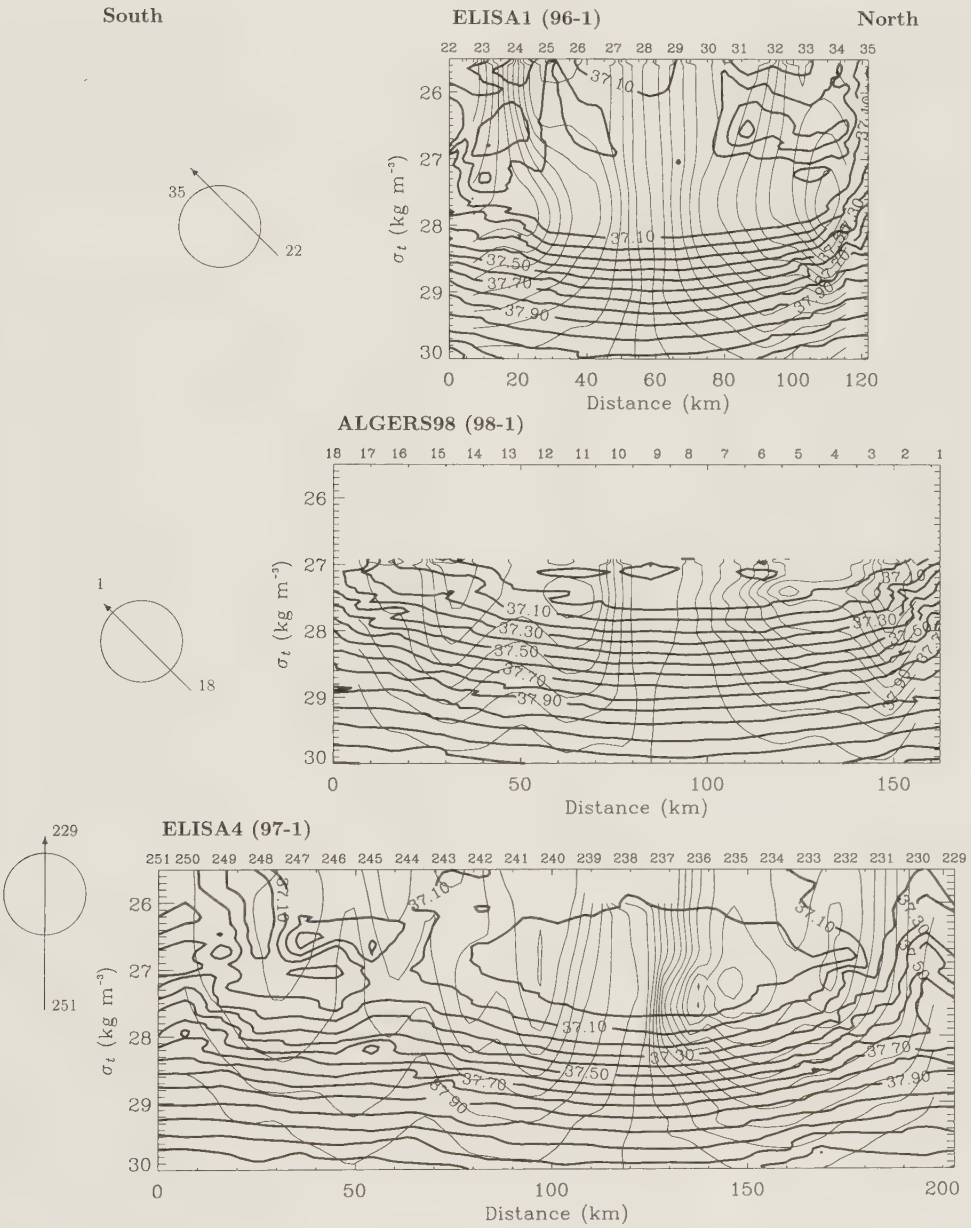


Fig. 9. Transects of salinity in isopycnal coordinates shown in Fig. 3 (thick lines) and gradient velocities (thin lines).

terms of the eigenvalues of the Lagrangian acceleration tensor (Hua and Klein, 1998). But the Okubo–Weiss separation is reasonably correct under the assumption of slowly evolving fields, which is not justified for general unsteady flows. However, the snapshots of W distribution in the three consecutive SLA maps of for 96-1, have shown that such assumption is reasonably valid. In the last two maps (Figs. 5B and C) the eddy structure and its surrounding does not change too

much. The core is qualitatively similar in intensity, shape and location, and the two marked areas of strong deformation around it are also similar and persistent.

A second point is clearly related with the data set used in the analysis. In our opinion this is the element that mainly constraints the results presented here. Data come from several sources and devices, characterized by different resolutions in both space and time. This limits the level of description we can reach. In SLA maps the calculation of W is straightforward and it has been previously corroborated as an efficient and precise way to identify large eddies in this particular region (Isern-Fontanet et al., 2003). Problems may come from the reliability of the SLA data when constructing these maps due to the filtering and interpolation techniques. However, several comparisons between SLA maps and AVHRR data, dynamic heights and drifting buoys indicate they do not differ too much in identifying the eddies studied here. Obviously the resolution of SLA maps does not allow resolving structures as submesoscale eddies and filaments, very important to understand mixing and stirring properties (e.g. Klein et al., 1998). These can be well resolved by other remote sensing techniques with higher resolution as infrared images. Here, we have shown a close connection between the temperature patterns in one SST image and the distribution of W derived from almost simultaneous SLA maps. Nevertheless, this result is quite qualitatively, at it has been shown recently, because SST patterns reflect, not only the stirring by the flow, but also the non-conservative character of the temperature (e.g. Abraham and Bowen, 2002).

While the eddy core boundary, defined as closed lines of $W = 0$, does not necessarily mean a line of maximum velocity (strictly valid for axisymmetric flow), the location of this boundary, as seen by altimetry, roughly corresponds to the location of a surface velocity extremum from CTD vertical sections. However, the vertical structure of velocities is much more complex, showing in some cases (97-1) the existence of several velocity extrema that complicates the proper identification of the core boundary (Fig. 7). The resolution of SLA maps is not enough to appreciate it because they have been filtered using a low-pass filter with cut-offs of the order of 40 km. Then, a smoothing and therefore reduction of geostrophic velocities and their gradients clearly affect the computation of W in the way of hiding the finer structure of the velocity field.

The use of salinity as a tracer has two main drawbacks. T-S diagrams of CTD data corresponding to the center of eddies show that the low salinity is constrained above the 27.5 kg/m^3 isopycnal, which limits the depth that can be analyzed. On the other hand, salinity is not a passive tracer since it is related to density and thus to geostrophic velocity. In any case, salinity distributions clearly indicate that the eddy core, although formed by quite homogeneous water (mainly Atlantic Water from the Algerian current, trapped when the AE was generated), is isolated from the circulation cell where water is mixed with ambient waters, then reaching characteristic values of the Mediterranean surface layer. Biological measurements in 96-1 during ELISA1 (Taupier-Letage et al., 2003) indicated the central zone was oligotrophic, as nutrients input in the surface layer associated with upwelling areas at the eddy periphery did not reach it. In 96-1 core, the quasi-homogeneous layer resulting from strong mixing during preceding spring(s) supports isolation. In the circulation cell, low-salinity filaments in the surface layer and less-mixed Levantine Intermediate Water patches support exchanges. Chlorophyll data and particulate matter distribution suggested different regimes of vertical motion associated with instability along the periphery of 96-1, while a comparison of the seston size spectra characteristics between samples in the core of 98-1 and others collected during MATER cruises in the basin indicated that the eddy was several months old with almost no exchange with surrounding waters (Rodríguez-Arias et al., 1998). It is interesting to comment the results of an experiment in which a chemical tracer was released within an eddy in the Atlantic North (Martin et al., 2001). The tracer cloud clearly evolve within the region between the center and the maximum velocity isotach. In terms of the separation given by W the tracer was released within the eddy core

region, and no exchanges were found with outer water mass.

The two distinct regions separated by W, eddy core and circulation cell, have different Lagrangian properties. The core is characterized by its impermeability to the exchange of particles with the circulation cell, which is therefore characterized by intense stirring with surrounding waters. In numerical experiments of two-dimensional turbulence it has been observed that particles in the circulation cell can be trapped for relatively long times (Elhmaïdi et al., 1993). The observation of drifters trajectories deployed during ALGERS98 cruise (Fig. 5E) shows how those deployed in a region close to the inner side of the core boundary followed closed trajectories around it. The four most external drifters moved initially very close and parallel but just arriving to the region of the stagnation point suffer an abrupt separation. Two of them escaped and the other two turned following a general path around the eddy. This separation of initially close drifters is characteristic of the critical streak line area and again consistent with the hyperbolic character of a persistent saddle point region. Recent studies on the Lagrangian properties of turbulent flows have shown that the stirring properties and their impact on tracer gradients are better studied from a Lagrangian point of view (Lapeyre et al., 2001; Haller, 2001). In general, the separation in hyperbolic and elliptic regions given by W is not equivalent with the presence of hyperbolic and elliptic Lagrangian trajectories. However, when the field is weakly unsteady a persistent region of hyperbolic character implies hyperbolic Lagrangian trajectories (Joseph and Legras, 2002).

6. Summary and conclusions

We have applied the Okubo–Weiss parameter, W, to analyze the spatial structure of three large mesoscale eddies in the Algerian basin. The spatial distribution of W allows distinguishing regions dominated by vorticity, eddy cores, from their surrounding regions where in general deformation dominates. Its implementation to sea level anomaly maps distinguishes quite precisely such regions for eddies in the Algerian basin. As seen in these maps, such eddies have very coherent near axisymmetrical core regions surrounded by circulation cells variable in both time and space.

The spatial distribution of such eddy cores, in terms of the Okubo–Weiss parameter, is very similar to the coherent eddies appearing in two-dimensional turbulence studies. Assuming that motion has axial symmetry the computation of W is equivalent to finding places where the radial velocity is maximum. This assumption allows connecting roughly the horizontal structure of W of an eddy with the vertical structure of velocities computed from CTD cross sections. Besides the vertical sections contain a much noisy and rich structure than the smooth horizontal structure from SLA maps, the agreement is reasonable. Since Algerian eddies have a well-defined vertical structure with a characteristic core of a rather homogeneous salinity distribution, salinity has been used to diagnose the validity of this qualitative argument. This has been tested in the isopycnal distribution of salinity for three AEs. Results have shown that for radii greater than the radial location of the velocity maximum, isohalines have a steep slope respect to isopycnals while in the inner part they are almost parallel. So at least qualitatively, there is a good agreement between the location of the maximum of velocities and the distribution of salinity gradients.

The spatial structure of W allows finding detailed information on the stirring properties of these eddies with implications on mixing. In the example shown, 96-1, the circulation cell contains persistent regions where deformation dominates over vorticity, high positive values of W. Such regions coincide with the position of stagnation points associated to critical streak lines. This helps to understand some thermal structures seen in SST images. In particular the stagnation regions are closely related with either a strong deformation of the thermal signature associated to the eddy and with the penetration of a filament of relative cold waters around the eddy. Such region determines the abrupt change of behavior of the Lagrangian motion of drifters deployed within the eddy. In this context, other magnitudes that characterize the Lagrangian dispersion, as Liapunov

exponents, are better to establish the local dispersion and to define kinematic boundaries or transport barriers associated to flow structures (Boffetta et al., 2001). These ideas have been recently applied in the atmosphere to study the mixing and exchanges through the limits of the Antarctic polar vortex (Joseph and Legras, 2002; Koh and Legras, 2002). Unfortunately, similar analysis cannot easily be implemented with our available data and are left for the future.

In summary, despite all the theoretical limitations the results presented here show the valuable character of the diagnose based on the Okubo–Weiss parameter, first to analyze and therefore understand sea surface patterns associated to marine eddies and second as a rather approximate method to diagnose the mixing and stirring properties from available Eulerian measurements.

Acknowledgements

This is a contribution to ELISA and ALGERS experiments and to the IMAGEN project funded by the Spanish R+D Plan (REN2001-0802-C02-02). Jordi Isern-Fontanet has been partially supported by contracts from GRAC (2FD97-0588) and IMAGEN projects. Altimetric maps for the period analyzed were elaborated and provided by CLS (Toulouse, France) under contract of the MATER project funded by the European Commission (MAS3-CT96-0051). ELISA1, ELISA4 and ALGERS98 were also partially funded by MATER. We would like to thank two anonymous reviewers for the comments that helped to improve the manuscript.

References

Abraham, E., Bowen, M., 2002. Chaotic stirring by a mesoscale surface-ocean flow. Chaos 12, 373–381.

AVISO User Handbook, 1996. Corrected Sea Surface Heights Products, 2.0 edition. AVI-NT-011-311-CN.

Ayoub, N., Le Traon, P., De Mey, P., 1998. A description of the Mediterranean surface variable circulation from combined ERS-1 and TOPEX/POSEIDON altimetric data. Journal of Marine Systems 18, 3–40.

Basdevant, C., Philipovitch, T., 1994. On the validity of th "Weiss criterion" in two-dimensional turbulence. Physica I 113, 17–30.

Benzohra, M., Millot, C., 1995a. Characteristics and circulation of the surface and intermediate water masses off Algeria Deep-Sea Research 42, 1803–1830.

Benzohra, M., Millot, C., 1995b. Hydrodynamics of an oper sea Algerian eddy. Deep-Sea Research 42, 1831–1847.

Boffetta, G., Lacorata, G., Redaelli, G., Vulpiani, A., 2001 Detecting barriers to transport, a review of different techniques. Physica D 159, 58–70.

Bouzinac, C., Vázquez, J., Font, J., 1998. CEOF analysis of ERS-1 and TOPEX/POSSEIDON combined altimetric data in the region of Algerian current. Journal of Geophysical Research 103, 8059–8071.

Bracco, A., LaCasce, J., Pasquero, C., Provenzale, A., 2000. The velocity distribution of barotropic turbulence. Physics of Fluids 12 (10), 2478–2488.

Brasseur, P., Beckers, J.M., Brankart, J., Shoenauen, R., 1996. Seasonal temperatures and salinity fields in the Mediterranean sea, climatological analysis of a historical data set. Deep-Sea Research 43, 159–192.

Dewar, W., Flierl, G., 1985. Particle trajectories and simple models of transport in coherent vortices. Dynamics of Atmospheres and Oceans 215–252.

Elhmaïdi, D., Provenzale, A., Babiano, A., 1993. Elementary topology of two-dimensional turbulence from a Lagrangian viewpoint and single-particle dispersion. Journal of Fluid Mechanics 257, 533–558.

Font, J., Millot, C., Salas, J., Julià, A., Chic, O., 1998. The drift of Modified Atlantic Water from the Alboran Sea to the eastern Mediterranean. Scientia Marina 62 (3), 211–216.

Font, J., Isern-Fontanet, J., Salas, J., 2004. Tracking a big anticyclonic eddy in the western Mediterranean Sea. Scientia Marina 68 (2), 331–342.

Fuda, J., Millot, C., Taupier-Letage, I., Send, U., Bocognano, J., 2000. XBT monitoring of a meridian section across the western Mediterranean Sea. Deep-Sea Research I 47, 2191 2218.

Gaspar, P., Ogor, F., Le Traon, P.Y., Zanife, O.Z., 1994. Estimating the sea state bias for TOPEX and POSEIDON altimeters from crossover differences. Journal of Geophysical Research 99 (C12), 24981–24994.

Haller, G., 2001. Lagrangian coherent structures and the rate of strain in two-dimensional turbulence. Physics of Fluids 13, 3365–3385.

Herbaut, C., Martel, F., Crépon, M., 1997. A sensitivity study of the general circulation of the Western Mediterranean sea. Part II, the response to atmospheric forcing. Journal of Physical Oceanography 27, 2126–2145.

Hua, B., Klein, P., 1998. An exact criterion for the stirring properties of nearly two-dimensional turbulence. Physica D 113, 98–110.

Isern-Fontanet, J., García-Ladona, E., Font, J., 2003. Identification of marine eddies from altimetry. Journal of Atmospheric and Oceanic Technology 20, 772–778.

Isern-Fontanet, J., García-Ladona, E., Font, J., 2004. The vortices of the Mediterranean sea: an altimetric perspective, submitted for publication.

Iudicone, D., Santoleri, R., Marullo, S., P., G., 1998. Sea level variability and surface eddy characteristics in the Mediterranean Sea from TOPEX/POSEIDON data. Journal of Geophysical Research 103(C2) 2995–3011.

Jeong, J., Hussain, F., 1995. On the identification of a vortex. Journal of Fluid Mechanics 285, 69–94.

Joseph, B., Legras, B., 2002. Relation between kinematic boundaries, stirring and barriers for the Antartic Polar Vortex. Journal of Atmospheric Sciences 59, 1198–1212.

Kamenkovich, I., Koshlyakov, A., Monin, A., 1986. Synoptic Eddies in the Ocean. D. Reidel Publishing Co., Dordrecht.

Klein, P., Tréguier, A., Hua, B., 1998. Quasigeostrophic stirring of thermohaline fronts. Journal of Marine Research 56, 589–612.

Koh, T., Legras, B., 2002. Hyperbolic lines and the stratospheric polar vortex. Chaos 12 (2), 382–394.

Lapeyre, G., Hua, B., Klein, P., 2001. Dynamics of active and passive tracers in a mesoscale turbulent flow field. Physics of Fluids 13, 251–264.

Larnicol, G., Le Traon, P., Ayoub, N., De Mey, P., 1995. Mean sea level and surface circulation variability of the Mediterranean Sea from 2 years of TOPEX/POSEIDON altimetry. Journal of Geophysical Research 100 (C12), 385–396.

Larnicol, G., Ayoub, N., Le Traon, P., 2002. Major changes in the Mediterranean sea level variability from 7 years of TOPEX/POSEIDON ERS-1/2 data. Journal of Marine Systems 33–34, 63–89.

Le Traon, P., Ogor, F., 1998. ERS-1/2 orbit improvement using TOPEX/POSEIDON: the 2cm challenge. Journal of Geophysical Research 103 (C4), 8045–8057.

Le Traon, P., Nadal, F., Ducet, N., 1998. An improved mapping method of multisatellite altimeter data. Journal of Atmospheric and Oceanic Technology 15, 522–534.

Martin, A., Richards, K., Law, C., Liddicoat, M., 2001. Horizontal dispersion within an anticyclonic mesoscale eddy. Deep-Sea Research Part II 48, 739–755.

McWilliams, J., 1984. The emergence of isolated coherent vortices in turbulent flow. Journal of Fluid Mechanics 146, 21–43.

Millot, C., 1985. Some features of the Algerian current. Journal of Geophysical Research 90 (C4), 7169–7176.

Millot, C., 1987. Circulation in the Western Mediterranean Sea. Oceanologica Acta 10, 143–149.

Millot, C., Benzohra, M., Taupier-Letage, I., 1997. Circulation off Algeria inferred from the médiprod-5 current meters. Deep-Sea Research 44, 1467–1495.

Obaton, D., Millot, C., Chabert D'Hières, G., Taupier-Letage, I., 2000. The Algerian current, comparisons between in situ and laboratory data sets. Deep-Sea Research I 47, 2159–2190.

Okubo, A., 1970. Horizontal dispersion of floatable particles in the vicinity of velocity singularities such as convergences. Deep-Sea Research 17, 445–454.

Olson, D., 1991. Rings in the ocean. Annual Review of Earth and Planetary Sciences 19, 283–311.

Ovchinikov, I., 1966. Circulation in the surface and intermediate layers of the Mediterranean. Oceanology 6, 48–59.

Pasquero, C., Provenzale, A., Babiano, A., 2001. Parametrization of dispersion in two-dimensional turbulence. Journal of Fluid Mechanics 439, 279–303.

Picco, P., 1990. Climatological atlas of the western Mediterranean. Technical Report, ENEA, La Spezia.

Provenzale, A., 1999. Transport by coherent barotropic vortices. Annual Review of Fluid Mechanics 31, 55–93.

Puillat, I., Taupier-Letage, I., Millot, C., 2002. Algerian eddies lifetime can near 3 years. Journal of Marine Systems 31, 245–259.

Robinson, A., 1983. Eddies in Marine Science. Springer, Berlin.

Rodríguez-Arias, M., Calafat, A., Canals, M., GRC Marine Geosciences, 1998. Differences in seston size spectra between the málaga upwelling system (Alboran sea) and anticyclonic stable eddy in the western Algerian basin. Technical Report, University of Barcelona.

Ruiz, S., Font, J., Emelianov, M., Isern-Fontanet, J., Millot, C., Salas, J., Taupier-Letage, I., 2002. Deep structure of an open sea eddy in the Algerian Basin. Journal of Marine Systems 33–34, 179–195.

Salas, J., 2003. Evolution of the open-sea eddy ALGERS'98 in the Algerian basin with Lagrangian trajectories and remote sensing observations. Journal of Marine Systems 43, 105–131.

Salas, J., García-Ladona, E., Font, J., 2001. Statistical analysis of the surface circulation in the Algerian Current using Lagrangian buoys. Journal of Marine Systems 29, 69–85.

Salas, J., Millot, C., Font, J., García-Ladona, E., 2002. Analysis of mesoscale phenomena in the Algerian basin from drifting buoys and infrared images. Deep-Sea Research I 49, 245–266.

Send, U., Font, J., Krahman, G., Millot, C., Rhein, M., Tintoré, J., 1999. Recent advances in studying the physical oceanography of the western Mediterranean sea. Progress in Oceanography 44, 37–64.

Taupier-Letage, I., Millot, C., 1988. Surface circulation in the Algerian basin during 1984. Oceanologica Acta 79–85.

Taupier-Letage, I., Puillat, I., Millot, C., Rimbault, P., 2003. Biological response to mesoscale eddies in the Algerian basin. Journal of Geophysical Research 108 (0), 10.1029/1999JC000117.

Tziperman, E., Malanotte-Rizzoli, P., 1991. The climatological seasonal circulation of the Mediterranean sea. Journal of Marine Research 49, 411–434.

Vazquez-Cuervo, J., Font, J., Martinez-Benjamin, J., 1996. Observations on the circulation in the Alboran sea using ERS1 altimetry and sea surface temperature data. Journal of Physical Oceanography 26 (8), 1426–1439.

Vignudelli, S., 1997. Potential use of ERS-1 an Topex/Posseidon altimeters for resolving oceanographic patterns in the Algerian Basin. Geophysical Research Letters 24 (14), 1787–1790.

Weiss, J., 1991. The dynamics of enstrophy transfer in two-dimensional hydrodynamics. Physica D 48, 273–294.

Wilkin, J., Morrow, R., 1994. Eddy kinetic energy and momentum flux in the Southern ocean, comparisons of a global eddy-resolving model with altimeter, drifter and current-meter data. Journal of Geophysical Research 99 (C4), 7903–7916.

ELSEVIER

Deep-Sea Research II 51 (2004) 3029–3042

DEEP-SEA RESEARCH
PART II

www.elsevier.com/locate/dsr2

An examination of the radiative and dissipative properties of deep ocean internal tides

Louis C. St. Laurent[a,*], Jonathan D. Nash[b]

[a]*Department of Oceanography, Florida State University, Tallahassee, Florida, USA*
[b]*College of Oceanic and Atmospheric Science, Oregon State University, Corvallis, Oregon, USA*

Abstract

In this study, the energy flux and energy dissipation of deep ocean internal tides are examined. Properties of the internal tide from two distinct generation regions are contrasted: the Mid-Atlantic Ridge (MAR) and the Hawaiian Ridge. Considerable differences are noted for the baroclinic energy flux, $\langle \mathbf{u}'p' \rangle$, radiated from each site. Radiation from the MAR is relatively rich in high modes, with an energy flux spectral peak at mode 5 and modes 10 and greater accounting for 40% of the total flux. In contrast, Hawaiian Ridge radiation is dominantly composed of modes 1 and 2, with modes 10 and greater accounting for less than 5% of the total flux. Depth integrated energy flux levels are $O(1)\,\mathrm{kW\,m^{-1}}$ at the MAR site, and $O(10)\mathrm{kW\,m^{-1}}$ at the Hawaiian Ridge. Despite these differences, observed turbulent dissipation rates at these sites are similar in magnitude and depth dependence. Decay scales, estimated as $L_\varepsilon = (\int_{-H}^{0} \mathbf{u}'p'\,\mathrm{d}z)/(\int_{-H}^{0} \rho_0 \varepsilon\,\mathrm{d}z)$, range from $O(100)$km to $O(1000)$ km. The mean decay scale based on the MAR data is 230 km, a factor of 3 smaller than at the Hawaiian Ridge site. We demonstrate that the dissipation level scales with the energy flux available in the high modes, which is comparable at both sites, rather than the total energy flux.
© 2004 Elsevier Ltd. All rights reserved.

1. Introduction

There is a renewed interest in internal tides, in part due to the improved observations of sea surface elevation by the Topex/Poseidon altimeter. These data have been used to constrain hydrodynamic models of the tides, and hence, produce accurate estimates of open-ocean barotropic tides (Egbert et al., 1994; Egbert, 1997). The residuals of such models have also been analyzed, and Egbert and Ray (2000, 2001) have interpreted patterns of implied barotropic to baroclinic tidal conversion. Specifically, these authors have identified a number of deep ocean regions where barotropic tidal energy is likely being transferred to internal tides.

While the global maps of Egbert and Ray provide insight into the spatial variations of the internal tides, they give no direct means of assessing in situ properties. Such studies require direct measurements from the ocean. While much work has focused on the internal tides generated

*Corresponding author.
E-mail address: lous@ocean.fsu.edu (L.C. St. Laurent).

0967-0645/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dsr2.2004.09.008

on continental shelves, canyons, straits, and fjords; only two prominent deep-ocean internal tide generation sites have been examined in dedicated field programs: the Mid-Atlantic Ridge site of the Brazil Basin Tracer Release Experiment (BBTRE), and the Hawaiian Ridge during the Hawaii Ocean Mixing Experiment (HOME). In the material that follows, we describe observations of internal tide energetics for each of these deep-ocean generation regions. While there are a number of publications describing various aspects of the internal tide at each of these sites, no previous attempt has been made to compare and contrast the measured properties. Such is the purpose of this article.

Numerous studies have addressed the internal tide generation problem. Original work by Cox and Sandstrom (1962), Baines (1973, 1982) and Bell (1975) provided the basis for later work. Further analytic formulations examined the initial transient wave response, finite depth effects, depth-varying stratification, spatial variations in topography and tidal forcing, and wave generation by finite amplitude topography of arbitrary steepness (Hibiya, 1986; St. Laurent and Garrett, 2002; Llewellyn Smith and Young, 2002, 2003; Balmforth et al., 2002; Khatiwala, 2003; St. Laurent et al., 2003; Petrelis et al., 2004). Numerical simulations of internal tide generation have also examined critical and supercritical generation phenomena, nonhydrostatic effects, instability, and mixing (Khatiwala, 2003; Legg and Adcroft, 2003; Legg, 2004a, b; Slinn and Levine, 2003). We neglect here references to the vast literature on coastal, hydraulically controlled, and nonlinear internal tide generation. Instead, the focus of this paper is on deep-ocean internal tides where linear dynamics and low Froude numbers generally characterize the wave generation. Moreover, a detailed discussion of internal tide generation is beyond the scope of this article. Instead, we focus on the properties of the internal tide energy flux that radiates away from two disparate generation sites.

In Section 2, we describe the internal tides and present background information for the BBTRE (Section 2.1) and HOME (Section 2.2) regions. A direct comparison of the M_2 internal tide at each site is presented in Section 2.3. Energy flux spectra for each region are presented in Section 3. Data describing internal tide dissipation and decay are presented in Section 4. Results are discussed in Section 5.

2. Estimates of baroclinic energy flux

2.1. The Mid-Atlantic Ridge of the Brazil Basin

The Mid-Atlantic Ridge (MAR) is a prominent topographic feature of the global ocean. The MAR, and mid-ocean ridges in general, are distinguished by their large-scale tectonic structure. The topography is comprised of many fractures, which contribute to a general roughness of the bathymetry. This roughness makes the MAR an efficient generation site for internal tides. Egbert and Ray (2000, 2001) estimate that 100 GW of barotropic to baroclinic tidal conversion occurs along the MAR, representing 10% of the global internal tide production.

Direct observations relating internal tides to mixing were found during the Brazil Basin Tracer Release Experiment (BBTRE; Polzin et al., 1997; Ledwell et al., 2000; St. Laurent et al., 2001). The study was conducted in the South Atlantic in a region spanning a network of fracture zones west of the Mid-Atlantic Ridge axis. In this region (Fig. 1), topographic relief varied by up to 1 km between the crests and floors of abyssal canyons, which had typical widths of 30–50 km. Full depth measurements of fine- and microstructure velocity were made at roughly 130 station locations above the rough topography. Tidal currents in this region have been examined using the TPXO model (Egbert, 1997), and the M_2 ellipse is dominant with an amplitude of $3\,\mathrm{cm\,s^{-1}}$.

Evidence implicating tides in driving deep mixing was presented by Ledwell et al. (2000) and St. Laurent et al. (2001). These authors present vertically integrated dissipation data that are modulated over the spring-neap tidal cycle with a lag of about a day. Maximum levels of dissipation, vertically integrated to 2000 m above the bottom, reach $3\,\mathrm{mW\,m^{-2}}$. As discussed by St. Laurent et al. (2001), internal waves generated by tidal flow over the topography are the only

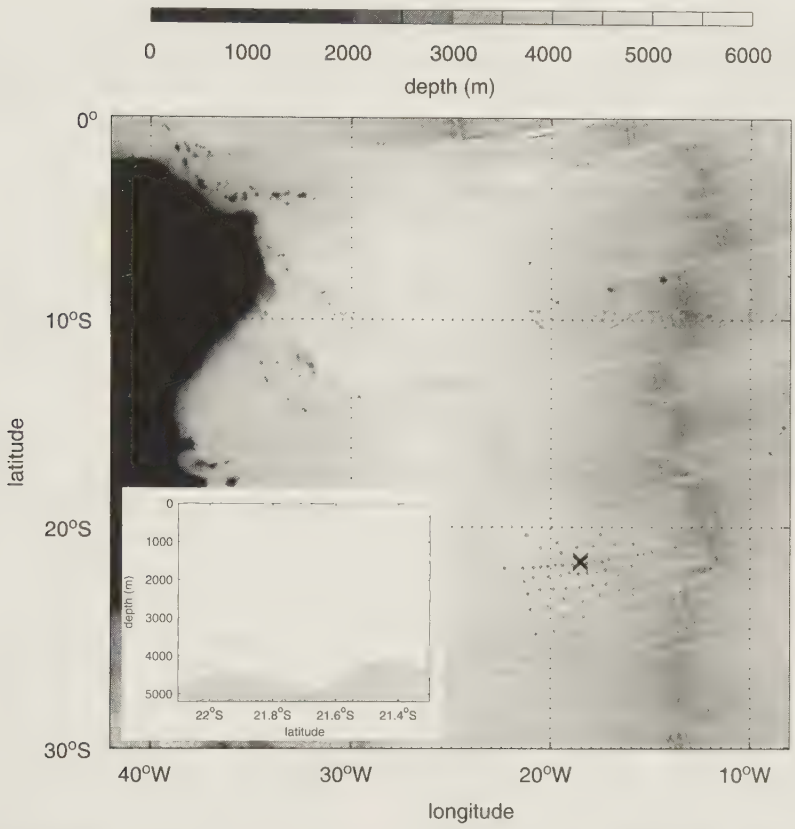


Fig. 1. Map showing the location of the Brazil Basin Tracer Release Experiment. The 1997 survey grid is shown along the flank of the MAR axis, which meanders between 12°W and 15°W. Bathymetry data from Smith and Sandwell (1997; version 8.2) were used in the figure. The location of a time series collected in 1996 at 18.5°W, 21.6°S is marked by the "x." The inset panel shows a north–south section of topography near the site of the time series.

plausible mechanism which can account for the observed energy level of the turbulence away from the bottom.

In the present analysis, we examine observations from two surveys conducted during BBTRE. The first survey consists of a time series of 17 profiles on the flank of a fracture zone canyon at 18.5°W, 21.6°S, collected over a 16-day period in 1996. The second survey consists of roughly 90 profiles spanning a nearly 1000 by 500 km region, collected over a 30-day period in 1997. The data presented here comes from the measurements of the High Resolution Profiler operated by the Woods Hole Oceanographic Institution (HRP; Schmitt et al., 1988). Data from both of these surveys are discussed by St. Laurent et al. (2001).

2.2. The Hawaiian Ridge

Like the MAR, the Hawaiian Ridge is a prominent topographic feature of the global ocean. Egbert and Ray (2000, 2001) estimate that over 20 GW of barotropic tidal energy is converted at Hawaii, much of it transferred to internal tides. Observations of internal tides were presented by Ray and Mitchum (1996, 1997), who used along-track Topex/Poseidon records to examine the surface manifestation of internal tides generated along the Hawaiian Islands. They found that both first and second baroclinic modes of the semi-diurnal (M_2) internal tides were present in the data, and that the signal of internal-tide propagation could be tracked at least 1000 km from the

Hawaiian Ridge. Ray and Mitchum (1997) assumed a two-layer stratification and estimated that 15 GW of semidiurnal internal tide energy radiates away from the Hawaiian Ridge in the first baroclinic mode. Ray and Cartwright (2001) later refined this to 6 GW using a model with more realistic stratification.

Direct observations of internal tides at the Hawaiian Ridge were made as part of the Hawaii Ocean Mixing Experiment (HOME; Pinkel et al., 2000). Both near- and far-field observational components were conducted, and fine- and microstructure measurements provided estimates of tidal energy flux and turbulent dissipation. Observations along the ridge were concentrated at a number of regions (Fig. 2), identified as having both strong (Nihoa Island, French Frigate Shoals and Kauai Channel) and weak (Necker Island) internal tide energy flux in numerical simulations (Merrifield et al., 2001). Preliminary HOME results have been reported by Rudnick et al. (2003). Here, we describe some properties of the baroclinic energy flux radiating from the Kaena Ridge of the Kauai Channel. Barotropic tides in this region are dominated by the M_2 component, with tidal velocities of $3\,\mathrm{cm\,s^{-1}}$ far from the ridge (similar to those at the BBTRE region of the MAR). Tidal speeds are much higher near the ridge, with currents of $10\,\mathrm{cm\,s^{-1}}$ in the 3000-m deep Kauai Channel increasing to $20\,\mathrm{cm\,s^{-1}}$ at water depths $<1000\,\mathrm{m}$.

We present time series observations from the 3000-m isobath of the Kauai Channel, collected in October 2000 during HOME (Kunze et al., 2004; Nash et al., 2004b). Measurements from eXpendable Current Profilers (XCPs; Sanford et al., 1993) and the Absolute Velocity Profiler (AVP; Sanford et al., 1978) were made by researchers at the University of Washington Applied Physical Laboratory. The 3000-m isobath site was occupied six times at 3-h intervals. A more comprehensive description of this survey and the treatment of finestructure data is given by Nash et al. (2004b).

2.3. Comparison of M_2 energy flux estimates

Here we examine full depth profiles of velocity and density finestructure. Specifically, we will estimate the baroclinic energy flux, $\mathbf{u}'p'$, where $\mathbf{u}'$ is the baroclinic velocity and p' is the baroclinic pressure anomaly. The baroclinic velocities are

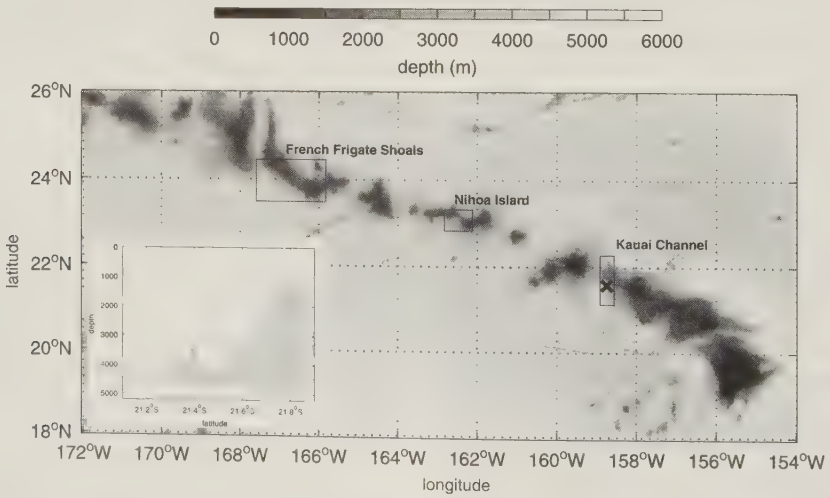


Fig. 2. Map showing several of the regions sampled during the Hawaii Ocean Mixing Experiment. French Frigate Shoals and the Kauai Channel are regions of intense baroclinic energy generation and were heavily sampled during 2000; Kauai Channel was the site of the intensive HOME nearfield experiment in 2002. Shading represents bathymetry data from Smith and Sandwell (1997; version 8.2). The location of a time series measured at the 3000-m isobath on the Kaena Ridge of the Kauai Channel (158.7°W, 21.6°N) is marked by the "x." The inset panel shows a north–south section of topography near the site of the time series.

computed directly from measured velocities: $u' = u - \bar{u}$ and $v' = v - \bar{v}$, where the barotropic components have been subtracted from each profile. The pressure anomaly was computed following the method of Kunze (Kunze et al., 2002; Nash et al., 2004c) using finestructure estimates of the buoyancy gradient N^2 and baroclinic displacement ξ,

$$p' = \rho_0 \int_z^0 N^2(z')\xi(z')\,dz' - \frac{\rho_0}{H}\int_{-H}^0\int_z^0 N^2(z')\xi(z')\,dz'dz. \quad (1)$$

The baroclinic displacements were estimated as $\xi = \zeta(\sigma;\mathbf{x},t) - \zeta(\bar{\sigma})$, where $\zeta(\sigma;\mathbf{x},t)$ are the instantaneous isopycnal depths observed at each station, and $\zeta(\bar{\sigma})$ are the averaged isopycnal depths. For the case of the BBTRE 1996 survey and the HOME 2000 survey, $\zeta(\bar{\sigma})$ was calculated using the time-mean isopycnal depths computed over the time series. For the BBTRE 1997 spatial survey, $\zeta(\bar{\sigma})$ is based on a least-square fit to each isopycnal surface of the form $\bar{\sigma}(x,y) = \varphi_{xx}x^2 + \varphi_x x + \varphi_{xy}xy + \varphi_y y + \varphi_{yy}y^2$, where the $\varphi_{(\cdot)}$ terms are the coefficients of the fit. A detailed description of this treatment of density data is given by St. Laurent et al. (2001).

Using the finestructure estimates of p' and $\mathbf{u}'$, profiles showing the instantaneous estimates $\mathbf{u}'p'$ can be computed for each of the surveys. However, only the time series data allow us to estimate the internal tide component of the baroclinic signal. Data from the BBTRE 1997 spatial survey is therefore not useful for this purpose, and will not be discussed further until Section 4.

The two time series surveys were examined using an identical analysis scheme. Vertical normal modes were computed for profiles of $\mathbf{u}'$ and p', solving the Sturm–Liouville equation numerically for the depth varying $N(z)$. The method is identical to that applied to mooring records (Alford, 2003), and results in a modal series of the form $p' \simeq \sum p_n F(z)$, where p_n are the modal amplitudes (for pressure anomaly) and $F(z)$ is the vertical structure function. A total of 25 modes were used to characterize the BBTRE profiles, while 14 were used for the HOME profiles. For each station location, a semidiurnal function of the form $\psi(t) = A\cos(\omega t) + B\sin(\omega t) + C$, was fit to the time series of normal mode amplitudes. This allows the extraction of the M_2 signal $\psi(t) - C$ with amplitude $\sqrt{A^2 + B^2}$ and phase $\phi = \arctan(A/B)$. The result is a set of profiles for baroclinic velocity and pressure anomaly that are roughly band-passed to the M_2 frequency and represented by a sum of vertical normal modes.

As described by Althaus et al. (2003) and Nash et al. (2004a), it is desirable to WKB scale the energy flux estimates as well as the depth coordinate. These transformations are given by $\hat{E}_f = (\mathbf{u}'p')\frac{N_0}{N(z)}$ and $\hat{z} = \int_z^0 \frac{N(z')}{N_0}\,dz'$, where $\hat{E}_f$ and $\hat{z}$ are the scaled energy flux and depth, respectively. For the BBTRE data, a thermocline value of $N_0 = 0.00524\,\mathrm{rad\,s^{-1}}$ was used as the reference stratification. For the HOME data, the mean stratification at the 3000-m isobath ($N_0 = 0.0029\mathrm{rad\,s^{-1}}$) was used. These scalings stretch the profiles to appear as if there was a depth constant stratification of N_0.

Fig. 3 presents profiles of the north–south component of the M_2 energy flux for the 17 profile BBTRE time series collected in 1996. These are instantaneous $v'p'$ estimates from single vertical profiles sampled at irregular intervals, so significant profile-to-profile variability is expected. In all profiles, there is a high-mode character to the energy flux. Noting the standard depth reference shown along the right-hand-side of the figure, the bottom-enhanced nature of the high-mode structure is evident. "Pulses" of $v'p'$ are most prominent in the depth range between 3000 m and the bottom, with each pulse occupying less than a 500-m vertical interval. Much weaker activity occurs in the upper 2000 m. These data suggest that the high-mode energy flux decays with a 1000–2000-m vertical scale as it radiates away from its generation site at the bottom. The mean energy flux is clearly to the north, which indicates that radiation is probably being emitted from the sloping southern side of the canyon (see Fig. 1). Energy flux levels in the mean profile are typically $1\,\mathrm{W\,m^{-2}}$, and the depth-integrated net energy flux is $1.0 \pm 0.1\,\mathrm{kW\,m^{-1}}$ in the north-northwest direction.

Fig. 4 shows profiles of the cross-ridge component of M_2 energy flux at the 3000-m isobath of

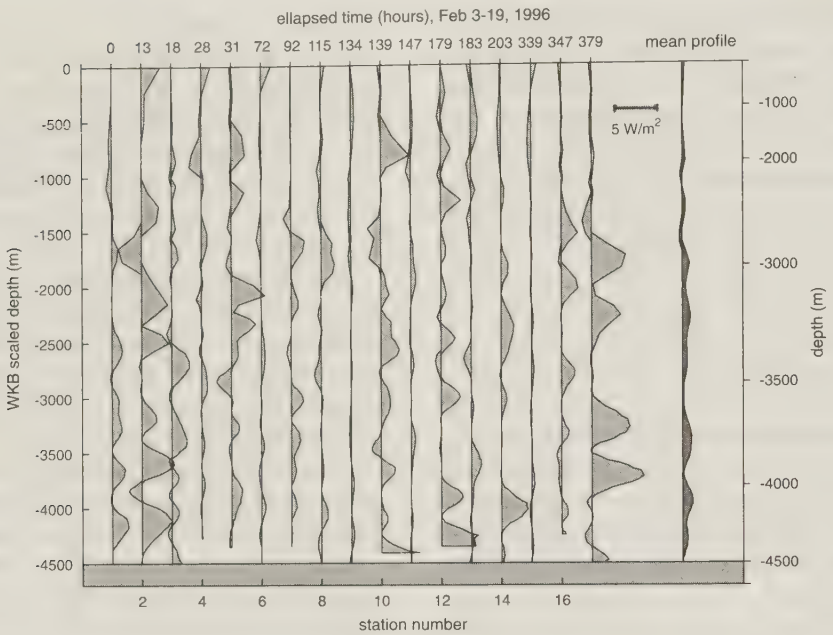


Fig. 3. A time series of instantaneous M_2 energy flux estimates $v'p'$ from 18.5°W, 21.6°S. This site is centered on a north–south slope between a 5000-m deep canyon to the south and a 4000-m deep ridge to the north. The energy flux estimates and the depth coordinate have been WKB scaled as described in the text, and the standard depth coordinate is given along the right side of the figure. Profiles were sampled at irregular intervals over a 16-day period, at times indicated along the upper axis. The mean profile is shown to the right. Energy flux plotted to the right is in the north direction.

the southern flank of the Kauai Channel. The instantaneous profiles were sampled at 3-h intervals, and a clear 6-h cycle is apparent in the time series, corresponding to the expected $\cos^2(\omega t)$ dependence of the energy flux. The baroclinic structure of energy flux is characterized by a low-mode structure, with much of the energy confined in two depth intervals, one in the upper portion of the water column and the other in the lower portion. Furthermore, each energy-flux bulge extends over a 1000-m unstretched depth scale. The mean energy flux profile for the Kauai Channel is shown to the right of Fig. 4. The flux is clearly directed south, away from the summit of the Kaena Ridge. Energy flux levels in the mean profile reach $50\,\mathrm{W\,m^{-2}}$, and the depth-integrated energy flux is $18 \pm 11\,\mathrm{kW\,m^{-1}}$.

The energy flux levels at the Kaena Ridge are $O(10)$ times larger than those measured above fracture zone topography in the Brazil Basin. This is not surprising, as the forcing of internal tides at Hawaii is stronger than at the MAR. This results from the greater topographic amplitudes (i.e., ridges $>3000\,\mathrm{m}$ high) associated with the 2000-km long Hawaii Ridge system. These intensify the barotropic currents as they are funneled through channels. In addition, the ridge topography protrudes through the thermocline, which leads to efficient baroclinic generation associated with regions of high N^2.

In addition to the disparity in energy flux magnitude, the baroclinic structure at each site is significantly different. Low modes dominate the flux at the Kaena Ridge, while high modes are clearly present at the BBTRE fracture zone site. This difference in baroclinic structure is fundamentally tied to differences in the topography at each internal tide generation site. In order to explore these details, energy flux spectra from generation models are now considered.

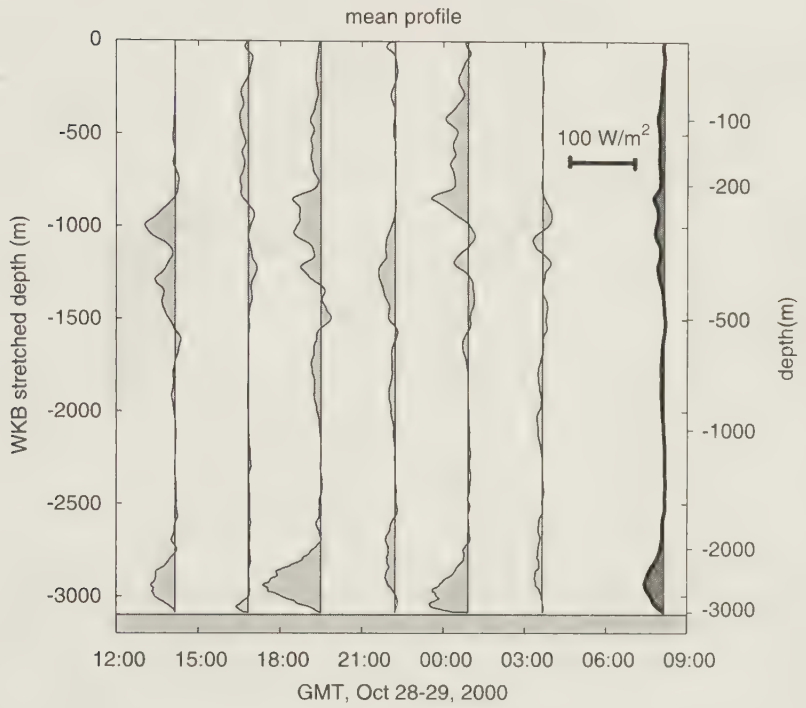


Fig. 4. Cross-ridge baroclinic energy flux density $v'p'$ radiating from Kaena Ridge, Kauai Channel. Energy flux estimates and the depth coordinate have been WKB scaled as described in the text. Instantaneous profiles from 6 occupations of the 3000-m isobath are shown. The samples were made at 3-h intervals. The mean profile is shown to the right. Energy flux plotted to the left is directed 37° west of south.

3. Energy flux spectra

An energy budget formulation for internal tides is presented by Khatiwala (2003). His formulation includes the relations between energy density, energy flux (lateral and vertical components), and the so-called barotropic to baroclinic "conversion rate." For the purpose of this article, the depth-integrated and time-averaged lateral energy flux, $\int_{-H}^{0}\langle u'p'\rangle\,\mathrm{d}z$, is regarded as equivalent to the conversion rate in the absence of dissipation. This is formally true when the tidal excursion, given by the ratio of the barotropic current amplitude to the tidal frequency, is much larger than the lateral scale of the topography. Under such conditions, the barotropic advection of baroclinic energy density is negligible. This is an excellent approximation for the internal tide generation conditions at the sites we consider in this study.

For the case of internal tides at the MAR, St. Laurent and Garrett (2002) used a linear model to examine the energy flux spectrum generated at the BBTRE site. Their calculations rely on a formulation similar to Bell (1975) and Llewellyn Smith and Young (2002), which compute energy flux based on the spectrum of bottom bathymetry. These formulations linearize the bottom boundary condition based on treating the topography as "subcritical" to internal tide generation. This refers to slope of the topography, $s = \nabla h$, being much less than the slope of the internal wave characteristics, $\alpha = \sqrt{(\omega^2 - f^2)/(N^2 - \omega^2)}$. St. Laurent and Garrett (2002) discuss in detail the applicability of the subcritical condition for MAR topography.

We present a spectral estimate for the semidiurnal internal tide energy flux at the BBTRE

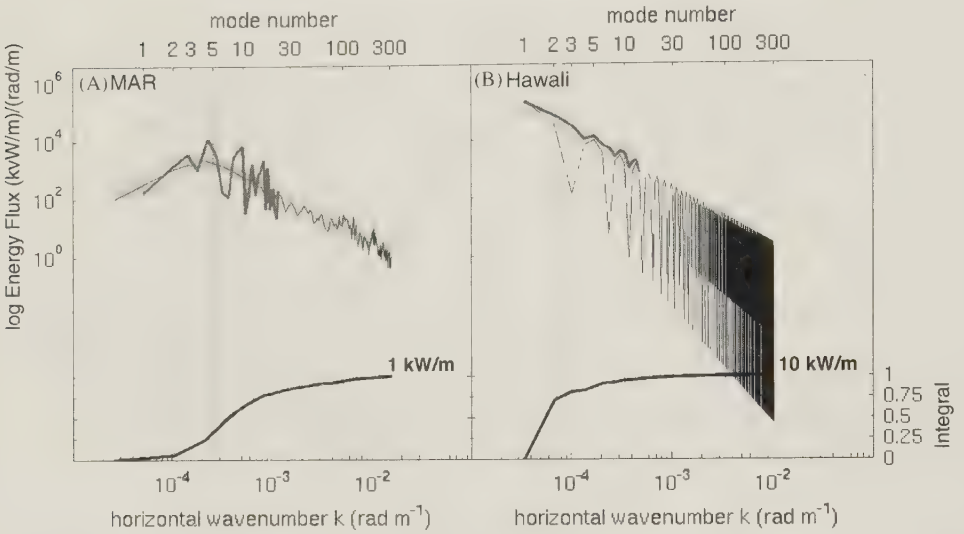


Fig. 5. (A) Spectrum for the depth-integrated average flux, $\int_{-H}^{0}\langle u'p'\rangle\,dz$, for subcritical internal tide generation at a Mid-Atlantic Ridge site. Multibeam bathymetry data were used to calculate the spectrum, and the standard error confidence band is shown. The wavenumber-integrated power level is roughly $1\,kW\,m^{-1}$. An observationally based estimate of the spectrum, $\int_{-H}^{0}\langle u_n'p_n'\rangle\,dz$, computed from 25 normal modes, is also shown with a 95% confidence interval. (B) Spectra for depth-integrated average flux, $\int_{-H}^{0}\langle u'p'\rangle\,dz$ for supercritical generation at a knife-edge ridge, used as a model for the Hawaiian Ridge. The wavenumber-integrated power level is $10\,kW\,m^{-1}$. An observationally based estimate of the energy flux spectrum $\int_{-H}^{0}\langle u_n'p_n'\rangle\,dz$, computed with 14 normal modes, is also shown with a 95% confidence interval. For both model estimates, the curve representing the cumulative integral of spectral power with wave number is shown.

region of the MAR in Fig. 5A. This estimate was derived from the MAR estimate of St. Laurent and Garrett (2002), who presented the time-averaged vertical component of the flux, $\langle w'p'\rangle$. Here, we have used their spectral model to calculate the horizontal energy flux, which we present in depth-integrated form, $\int_{-H}^{0}\langle u'p'\rangle dz$. As shown by St. Laurent and Garrett (2002), the MAR energy flux spectrum peaks near mode 5, with modes 1–10 accounting for 60% of the total energy flux. The integrated spectrum yields an estimate of roughly $1\,kW\,m^{-1}$. A spectral estimate of the energy flux was also made using the data comprising the mean profile shown in Fig. 3 (appearing on the right). This observationally based estimate is derived from the vertical-mode amplitudes and phases. The spectrum was computed as the energy flux per mode: $|u_n||p_n|\cos(\theta)\int_{-H}^{0}F^2(z)\,dz$, where $\theta=\phi_u-\phi_p$ is the phase difference between u_n and p_n, and $F(z)$ is the vertical structure function of the nth mode. Confidence intervals on the spectrum were computed assuming a χ^2 cumulative density for error on the mean with 17 degrees of freedom (the number of profiles contributing to the average). The integrated power in the estimated spectrum is 2 kW/m, a factor of 2 greater than the model spectrum. However, this difference is not statistically significant considering the 95% confidence interval of the estimates. Furthermore, the model spectrum is based on multibeam bathymetry data from a slightly different region to the south of the BBTRE region, where bathymetric statistics and tidal forcing may be slightly different. Overall, there is good qualitative agreement between the shapes of the observationally and model derived-spectra, which both show peak energy levels at roughly mode 5.

Energy flux estimates have also been considered for "supercritical" topographies. In particular, St. Laurent et al. (2003) and Llewellyn Smith and Young (2003) have formulated models for calculating the energy flux generated by an abrupt tall ridge. These models are based on a knife-edge ridge; i.e., a ridge with no width along the

direction of the tidal flow. The knife-edge ridge with a height of $\frac{3}{4}$ the water depth was used to approximate the Hawaiian Ridge, and the resulting estimate of energy flux is shown in Fig. 5B. The knife-edge model spectrum is accompanied by an estimate based on the observed M_2 energy flux at the 3000-m isobath of the Kauai Channel, using the mean profile shown in Fig. 4. Wave modes for the knife-edge ridge were calculated for a water depth of 3000 m using $N = 0.004\,s^{-1}$ and $f = 5.5 \times 10^{-5}\,s^{-1}$. As shown in Fig. 5B, both the knife-edge and observationally based spectra are "red." Furthermore, the spectra show excellent correspondence in modal content. Modes 1–10 account for more than 95% of both the modeled and observed energy flux; in fact, modes 1 and 2 account for 88% of the baroclinic energy flux in both the observationally based and model estimate. The generation of low mode internal tides is clearly favored by both rough MAR topography and abrupt Hawaii-like ridges alike. However, the relative proportion of high modes to low modes is significantly higher at the Mid-Atlantic Ridge site.

The shear (u_z) of the generated internal tides is also of interest, as the Richardson number ($Ri = N^2/u_z^2$) is a crucial stability parameter for baroclinic flow. St. Laurent and Garrett (2002) examined the shear of internal tides generated at MAR topography. They computed the cumulative integral of shear variance as a function of vertical wavenumber ($u_z^2(m)$), and then defined $Ri(m) = N^2/\langle u_z^2(m)\rangle$. This provides a depth-average estimate of the shear stability of the internal tide, as opposed to an estimate of stability at any given depth. They found that, in the absence of background sheared currents, the Richardson number is large for the wavelengths that dominate the energy flux (Fig. 6A). St. Laurent and Garrett (2002) therefore conclude that shear instability is not a primary influence on the low mode energy flux generated at subcritical topography. The Richardson number for the knife-edge ridge discussed above is also considered here (Fig. 6B). As is the case for the MAR model, the Richardson number, based on mean-square shear for the knife model, is characterized by $Ri > 1$. We note that in

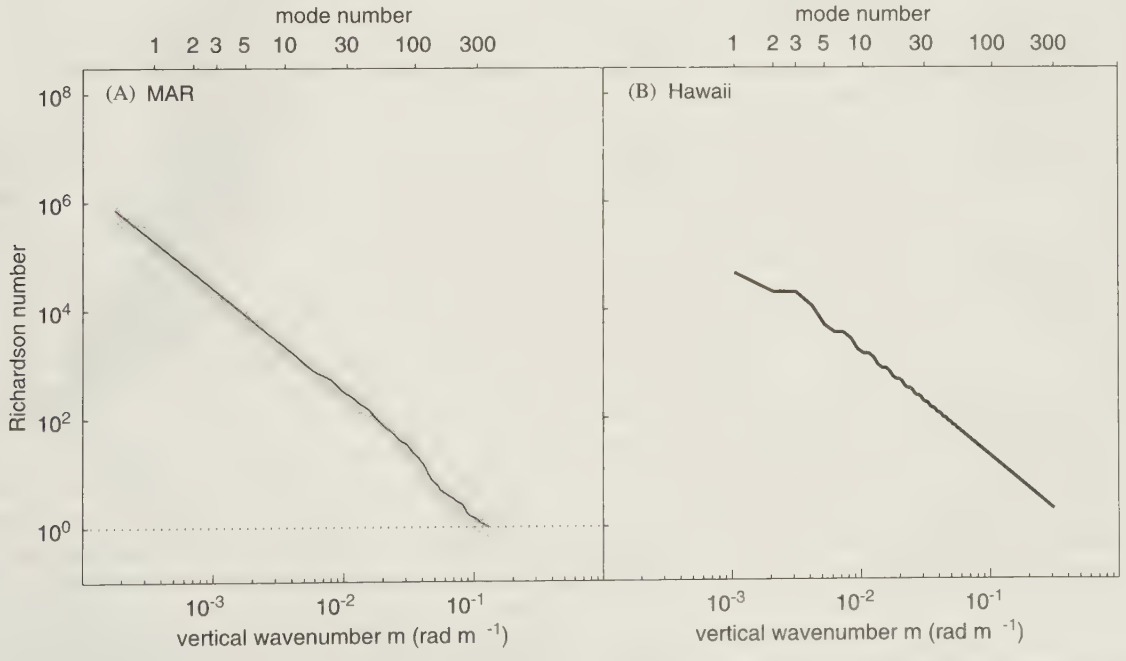


Fig. 6. The Richardson number estimated by integrating the cumulative shear variance of increasing modes for internal tides generated at (A) a mid-ocean ridge and (B) a knife-edge model approximation of the Hawaiian Ridge. In both panels, the first 300 baroclinic modes are shown. A Richardson number of unity is indicated for reference.

both cases, the Richardson numbers will drop below unity when very high modes are accounted for in the shear variance. In fact, for the case of supercritical wave generation, inviscid theories, like that of St. Laurent et al. (2003) for knife-edge generation, give infinite integrated shear variance. These extremely high modes manifest their shear in a narrow beam, as are observed at supercritical generation sites. As shown by Lien and Gregg (2001), large amounts of dissipation can occur along the path of the beam. However, it seems unlikely that the energy flux carried by the lowest modes is influenced by the low Richardson numbers associated with beam-like structures.

Subtle but important differences in *Ri* are evident between the HOME and BBTRE data shown in Fig. 6. Despite the low-mode internal tide being 10 times more energetic at the Hawaiian Ridge, its energy flux spectrum is redder so that the energy at high-wavenumbers is less than the MAR. This is illustrated by the fact that the spectral estimate of *Ri* from modes 1–300 is a factor of 3 larger at the Hawaiian Ridge than at the MAR. Hence, internal tides with 10-m vertical wavelength at the MAR may be susceptible to shear instability, while those same wavelengths at the HOME site may be stable, simply due to differences in the shape of the energy spectrum. Since turbulent overturns at the Hawaiian Ridge are observed to be many hundred meters in vertical extent (Slinn and Levine, 2003; Klymak et al., 2004), it is possible that the mechanism for dissipation at HOME is not the direct generation at M_2 frequencies, as was suspected at the MAR. Instead, different mechanisms may be in play at the Hawaiian Ridge; parametric subharmonic instability has been suggested as one such possibility to move variance directly from stable M_2 waves to high wavenumber, lower-frequency motions that may be susceptible to shear instability (i.e., MacKinnon and Winters, 2003).

We again emphasize that the spectral estimate of *Ri* is a depth-averaged quantity. It underestimates the amount of localized shear when a superposition of waves constructively interfere or are spatially phased-locked and form beam-like structures. This is likely near the Hawaiian Ridge (where phase-locking is observed), but has not been quantified.

4. Internal tide dissipation and decay

Direct measurements of velocity microstructure taken during BBTRE and HOME allow us to examine profiles of the energy dissipation rate (ε) at each internal tide generation site. Average profiles are shown in Fig. 7. Dissipation rate data from BBTRE have been previously discussed by Ledwell et al. (2000) and St. Laurent et al. (2001), and dissipation rate estimates from HOME are described in Rudnick et al. (2003) and Klymak et al. (2004). Despite the fact that energy flux levels are stronger at the Hawaiian Ridge, the average dissipation profiles from HOME and BBTRE are

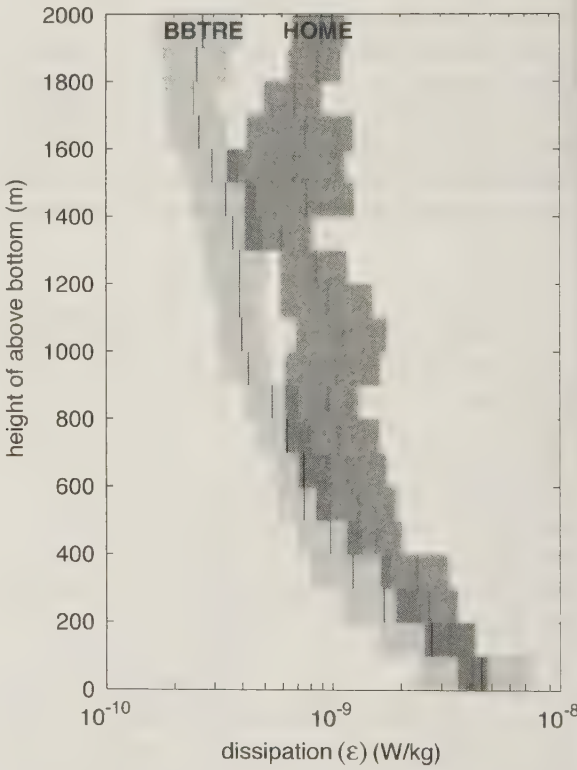


Fig. 7. Profiles of the turbulent dissipation rate ε from BBTRE and HOME observations. In both cases, dissipation rates were averaged into 100-m vertical bins and smoothed by a 3-bin running average. The 95% confidence band is shown for all estimates. The BBTRE profile is an average of 129 profiles from surveys in 1996 and 1997 spanning a 1000 km by 500 km region. The HOME profile is an average over 70 profiles spanning sites along the 3000-m isobath of the Hawaiian Ridge from French Frigate Shoals to Kauai Channel.

similar in both magnitude and depth dependence up to heights of 800 m above the bottom. At each site, the dissipation rate is maximum just above the topography, and the profiles decay with height with an e-fold scale of roughly 300–500 m. Above 800 m, the average profile from the Hawaiian Ridge shows dissipation values significantly larger than those observed at the MAR. We note that this difference is partially due to the buoyancy scaling of ε, as discussed by Polzin et al. (1995). Heights above 800 m at the MAR extend into the weakly stratified abyssal interior, while considerably more stratification exists at these heights along the Hawaiian Ridge.

Simultaneous observations of fine- and microstructure made during BBTRE and HOME permit estimation of a spatial decay scale for the radiated internal tides. For each profile, the depth-integrated energy flux was normalized by the depth-integrated turbulent energy dissipation to yield a "decay length" of the radiated tide, given by

$$L_\varepsilon \simeq \frac{\int \langle \mathbf{u}'p' \rangle \, \mathrm{d}z}{\int \rho \varepsilon \, \mathrm{d}z}, \tag{2}$$

as shown in Fig. 8. In (2), the vertical integrals were computed over the depth range between 100 m and the bottom (i.e., neglecting the surface-influenced 100 m). This decay scale represents the e-folding scale for an internal wave assuming the dissipation rate is constant along the path of propagation and negligible radial dispersion. This is obviously an approximate assumption, as the dissipation rate clearly varies spatially. Most likely, L_ε is an underestimate of the decay scale, since ε decays with distance from its energy source (Klymak et al., 2004). Nevertheless, the estimates presented here provide an order of magnitude scale for the decay of internal tide energy as it radiates away from regions of generation. Decay scales at the MAR (panel a) are less than those observed at the Hawaiian Ridge (panel b). The mean decay scale for the BBTRE data is 230 ± 70 km, and 750 ± 240 km for the HOME data. The difference in these mean estimates is significant at the 95% confidence level. Additionally, the minimum and maximum L_ε for BBTRE are 77 and 460 km, while these are 150 and 2200 km for

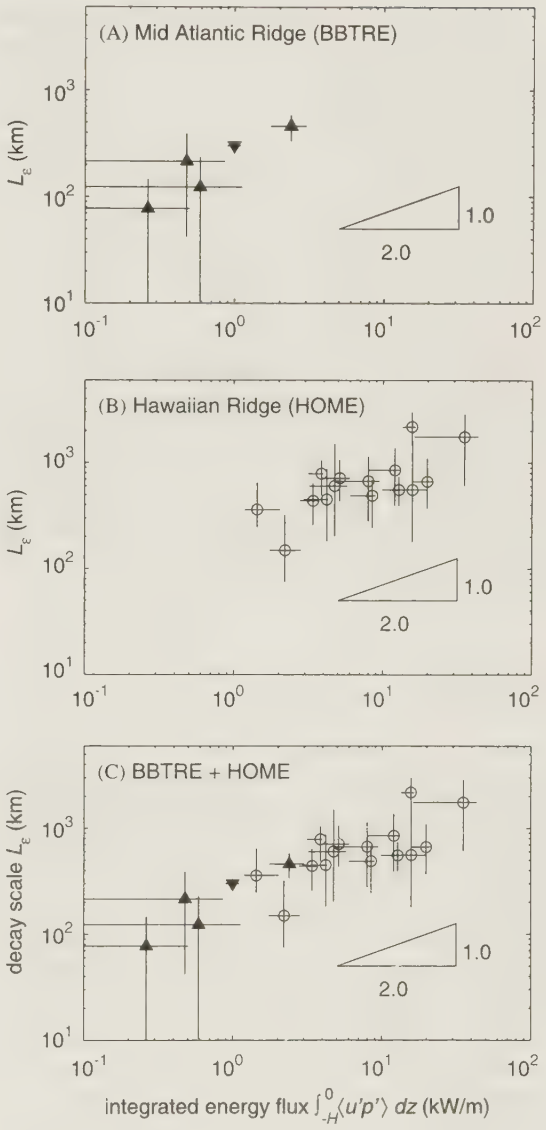


Fig. 8. Spatial decay scale L_ε as a function of depth-integrated energy flux. (A) Estimates from BBTRE 1996 (downward triangle) and BBTRE 1997 (upward triangles). Roughly 15 profiles were averaged for each estimate, and 95% confidence intervals are shown. (B) Estimates from the 3000-m isobath at the Hawaiian Ridge; each symbol represents a >5 profile average from each of 16 different station occupations. (C) Combined estimates from BBTRE and HOME. Error bars represent 95% bootstrap confidence intervals.

HOME. The $O(100)$km decay scale for the MAR has not been previously documented. However, the scale is consistent with the abrupt decrease in

turbulence levels away from the rough topography of the MAR (Polzin et al., 1997). The O(1000)km decay scales estimated at Hawaii are consistent with the correlation scales documented by Ray and Mitchum (1996) for the lowest mode internal tide radiated from the Hawaiian Ridge, although some of the decay perceived from the satellite measurements may result from loss of coherence of the M_2 tide as it passes through mesoscale features (Rainville, pers. comm.).

Fig. 8 shows that the largest decay scales are associated with the largest energy fluxes. We associate large energy flux with abundant low mode energy, which is able to radiate over large distances. We postulate the presence of dissipative high modes is only weakly related to the amount of low mode energy flux generation, so that dissipation occurs in higher proportion at sites of weak generation.

5. Discussion

We have presented estimates of baroclinic energy flux from two very different internal tide generation regions. The Mid-Atlantic Ridge site studied as part of BBTRE provides an example of internal tide generation by deep rough topography. In contrast, the Hawaiian Ridge sites studied during HOME provide examples of internal tide generation at abrupt high-amplitude topography. These two generation-regions provide examples from significantly different baroclinic wave forcing regimes of the deep ocean. As discussed by St. Laurent and Garrett (2002), the MAR is characterized by subcritical slopes, in that the radiated path of the internal tide is steeper than the large-scale slope of the bottom topography. This contrasts the Hawaiian Ridge, where internal tides propagate along ray paths less steep than the bottom. Nash et al. (2004b) show that the observed beams of M_2 energy at HOME are consistent with localized generation over the steep supercritical flanks (i.e., regions of strong barotropically induced vertical displacements; Baines, 1982) and subsequent propagation along internal tide ray paths. The beam-like nature near the ridge crest evolves into a more modal character by the 3000-m isobath, the site of dissipation measurements presented here.

Estimates of energy flux from the BBTRE site show considerable high mode character (Fig. 3). This is consistent with the modeled spectrum of energy flux (Fig. 5A), which shows an energy flux maximum at mode 5. Such high mode generation is clearly linked to the scales of the dominant fracture zone topography. The horizontal wavelength of the mode-5 M_2 internal tide is 30 km, roughly matching the characteristic wavelength of the canyon-crest system. Finer topographic scales are present in the form of abyssal hills in the canyons and on the ridge crests. These lead directly to high wavenumber generation, as discussed by Polzin (2004).

In contrast, the energy flux from the Hawaiian Ridge originates as low modes because of the topographical scale. The horizontal span of the Kaena Ridge appears to preferentially favor generation into modes 1 and 2, which produces beams along the summit of the ridge (Nash et al., 2004b) that quickly evolve into low modes by the 3000-m isobath (Fig. 4). As shown by the spectral estimate (Fig. 5B), the energy flux spectrum is "red," with most of the energy in modes 1 and 2. St. Laurent et al. (2003) suggest that low-mode energy flux is a general feature of internal tides generated by all abrupt tall ridges in the ocean. While high wavenumbers are generated in order to form the beam-like baroclinic structure, they represent only a small fraction of the energy flux and likely dissipate more quickly than their low-mode counterparts.

Despite the considerable differences in the radiated energy flux at these sites, the rate of energy dissipation of the internal tide is similar. Maximum dissipation levels of $\varepsilon = 10^{-9} - 10^{-8}\,\mathrm{W\,kg^{-1}}$ occur along the bottom in both regions (Fig. 7). Additionally, the height dependence of the dissipation profiles are of similar scale, particular within 800 m of the bottom. Decay scales, as estimated by normalizing the integrated energy flux by the integrated dissipation, range from O(100)km to O(1000)km. The mean estimate of L_ε is a factor of 3 less for the MAR. This is consistent with the higher-mode content of the MAR energy flux, as high modes are

more likely to dissipate close to their generation site (St. Laurent and Garrett, 2002).

In conclusion, the significant differences in internal tide generation at the MAR and Hawaiian Ridge do not lead to significant differences in the magnitude of the turbulent energy dissipation. If this result is general, there are interesting implications concerning the underlying physics of dissipation. It would seem that the overall level of dissipation is not sensitive to the total energy flux. Instead, we propose that the dissipation level scales only with the energy available in the high modes. A rough analysis follows from considering the energy flux in modes >10 for each site. For the MAR site, modes >10 comprise 40% of the 1–2 kW m^{-1} of depth integrated flux (Fig. 5A), thus accounting for 0.4–0.8 kW m^{-1}. Similarly, modes >10 at the Hawaiian Ridge comprise 5% of the 10 kW m^{-1} of integrated flux (Fig. 5B), accounting for 0.5 kW m^{-1}. Thus, while the overall energy flux levels differ, the high-mode energy level is comparable at the two sites. The high-mode energy density may thus provide a better scaling for the turbulent energy dissipation rate, as suggested by Klymak et al. (2004).

As argued by St. Laurent and Garrett (2002), energy flux in high modes will dissipate locally, as their group speed is very small and their wave–wave interaction rates are large. This leaves the low modes to radiate their energy away, apparently to considerable distances. Where and how the low mode energy dissipates remains an open question.

Acknowledgements

Data presented in this report were collected by a number of investigators. We thank Tom Sanford, Eric Kunze and Craig Lee for their efforts in the collection and interpretation of Hawaii Ocean Mixing Experiment data presented in this report. Additionally, we are grateful to John Toole, Ray Schmitt, and Kurt Polzin for their efforts associated with the BBTRE program and data. We thank Eric Kunze, Jen Mackinnon, Rob Pinkel, and two anonymous reviewers for their helpful comments. This research has been funded by grants from NSF and ONR.

References

Alford, M.H., 2003. Redistribution of energy available for ocean mixing by long-range internal-wave propagation. Nature 423, 159–163.

Althaus, A.M., Kunze, E., Sanford, T.B., 2003. Internal tide radiation from Mendocino Escarpment. Journal of Physical Oceanography 33, 1510–1527.

Baines, P.G., 1973. The generation of internal tides by flat-bump topography. Deep-Sea Research 20, 179–205.

Baines, P.G., 1982. On internal tide generation models. Deep-Sea Research 29, 307–338.

Balmforth, N.J., Ierley, G.R., Young, W.R., 2002. Tidal conversion by subcritical topography. Journal of Physical Oceanography 32, 2900–2914.

Bell, T.H., 1975. Lee waves in stratified flows with simple harmonic time dependence. Journal of Fluid Mechanics 67, 705–722.

Cox, C.S., Sandstrom, H., 1962. Coupling of surface and internal waves in water of variable depth. Journal of the Oceanographic Society of Japan, 20th Anniversary Volume, pp. 499–513.

Egbert, G.D., 1997. Tidal data inversion: interpolation and inference. Progress in Oceanography 40, 81–108.

Egbert, G.D., Ray, R.D., 2000. Significant dissipation of tidal energy in the deep ocean inferred from satellite altimeter data. Nature 405, 775–778.

Egbert, G.D., Ray, R.D., 2001. Estimates of M2 tidal energy dissipation from TOPEX/POSEIDON altimeter data. Journal of Geophysical Research 106, 22475–22502.

Egbert, G.D., Bennett, A.F., Foreman, M.G.G., 1994. TOPEX/POSEIDON tides estimated using a global inverse model. Journal of Geophysical Research 99, 24821–24852.

Hibiya, T., 1986. Generation mechanism of internal waves by tidal flow over a sill. Journal of Geophysical Research 91, 7696–7708.

Khatiwala, S., 2003. Generation of internal tides in an ocean of finite depth: analytical and numerical calculations. Deep-Sea Research 50, 3–21.

Klymak, J.M., Moum, J.N., Nash, J.D., Kunze, E., Girton, J.B., Carter, G.S., Lee, C.M., Sanford, T.B., Gregg, M.C., 2004. An estimate of energy lost to turbulence at the Hawaiian Ridge. Journal of Physical Oceanography, submitted.

Kunze, E., Rosenfeld, L.K., Carter, G.S., Gregg, M.C., 2002. Internal waves in the Monterey submarine canyon. Journal of Physical Oceanography 32, 1890–1913.

Kunze, E., Sanford, T.B., Lee, C.M., Nash, J.D., Merrifield, M.A., Holloway, P., 2004. A survey of internal tides and turbulence along the Hawaiian Ridge with model comparisons. Journal of Physical Oceanography, in preparation.

Ledwell, J.R., Montgomery, E.T., Polzin, K.L., St. Laurent, L.C., Schmitt, R.W., Toole, J.M., 2000. Mixing over rough topography in the Brazil Basin. Nature 403, 179–182.

Legg, S., 2004a. Internal tides generated on a corrugated continental slope. Part I: cross-slope barotropic forcing. Journal of Physical Oceanography 34, 156–173.

Legg, S., 2004b. Internal tides generated on a corrugated continental slope. Part II: along slope barotropic forcing. Journal of Physical Oceanography 34, 1824–1838.

Legg, S., Adcroft, A., 2003. Internal wave breaking at concave and convex continental slopes. Journal of Physical Oceanography 33, 2224–2246.

Lien, R.-C., Gregg, M.C., 2001. Observations of turbulence in a tidal beam and across a coastal ridge. Journal of Geophysical Research 106, 4575–4592.

Llewellyn Smith, S.G., Young, W.R., 2002. Conversion of the barotropic tide. Journal of Physical Oceanography 32, 1554–1566.

Llewellyn Smith, S.G., Young, W.R., 2003. Tidal conversion at a very steep ridge. Journal of Fluid Mechanics 495, 175–191.

MacKinnon, J.A., Winters, K.B., 2003. Spectral evolution of bottom-forced internal waves. Proceedings of the 13th 'Aha Huliko' a Hawaiian Winter Workshop, pp. 73–83.

Merrifield, M.A., Holloway, P.E., Shaun Johnston, T.M., 2001. Internal tide generation at the Hawaiian Ridge. Geophysical Research Letters 28, 559–562.

Nash, J.D., Kunze, E., Toole, J.M., Schmitt, R.W., 2004a. Internal tide reflection and turbulent mixing on the continental slope. Journal of Physical Oceanography 34 (5), 1117–1134.

Nash, J.D., Kunze, E., Sanford, T.B., Lee, C.M., 2004b. Structure of the baroclinic tide generated at a ridge. Journal of Physical Oceanography, submitted.

Nash, J.D., Alford, M.H., Kunze, E., 2004c. On estimating internal-wave energy fluxes in the ocean. Journal of Atmospheric Oceanic Technology, submitted.

Petrelis, F., Llewellyn Smith, S.G., Young, W.R., 2004. Tidal conversion at a submarine ridge. Journal of Physical Oceanography, in press.

Pinkel, R., Munk, W., Worcester, P., et al., 2000. Ocean mixing studied near Hawaiian Ridge. EOS Transactions of the American Geophysical Union 81, pp. 545,553.

Polzin, K.L., 2004. Idealized solutions for the energy balance of the finescale internal wave field. Journal of Physical Oceanography 34, 231–246 doi:10.1175/1520-0485(2004)034.

Polzin, K.L., Toole, J.M., Schmitt, R.W., 1995. Finescale parameterizations of turbulent dissipation. Journal of Physical Oceanography 25, 306–328.

Polzin, K.L., Toole, J.M., Ledwell, J.R., Schmitt, R.W., 1997. Spatial variability of turbulent mixing in the abyssal ocean. Science 276, 93–96.

Ray, R.D., Cartwright, D.E., 2001. Estimates of internal tide energy fluxes from Topex/Poseidon altimetry: Central North Pacific. Geophysical Research Letters 28, 1259–1263.

Ray, R., Mitchum, G.T., 1996. Surface manifestation of internal tides generated near Hawaii. Geophysical Research Letters 23, 2101–2104.

Ray, R., Mitchum, G.T., 1997. Surface manifestation of internal tides in the deep ocean: observations from altimetry and island gauges. Progress in Oceanography 40, 135–162.

Rudnick, D.L., Boyd, T.J., Brainard, R.E., et al., 2003. From tides to mixing along the Hawaiian Ridge. Science 301, 355–357.

Sanford, T.B., Drever, R.G., Dunlap, J.H., 1978. A velocity profiler based on the principles of geomagnetic induction. Deep-Sea Research 25, 183–210.

Sanford, T.B., D'Asaro, E.A., Kunze, E.L., Dunlap, J.H., Drever, R.G., Kennelly, M.A., Prater, M.D. Horgan, M.S., 1993. An XCP user's guide and reference manual. Applied Physics Laboratory Technical Report. APL-UW TR9309, University of Washington, Seattle, WA.

Schmitt, R.W., Toole, J.M., Koehler, R.L., Mellinger, E.C., Doherty, K.W., 1988. The development of a fine- and microstructure profiler. Journal of Atmospheric and Oceanic Technology 5, 484–500.

Slinn, D.N., Levine, M.D., 2003. Modeling internal tides and mixing over ocean ridges. Proceedings of the 13th 'Aha Huliko' a Hawaiian Winter Workshop, pp. 59–68.

St. Laurent, L.C., Garrett, C., 2002. The role of internal tides in mixing the deep ocean. Journal of Physical Oceanography 32, 2882–2899.

St. Laurent, L.C., Toole, J.M., Schmitt, R.W., 2001. Buoyancy forcing by turbulence above rough topography in the abyssal Brazil Basin. Journal of Physical Oceanography 31, 3476–3495.

St. Laurent, L.C., Simmons, H.L., Jayne, S.R., 2002. Estimating tidally driven mixing in the deep ocean. Geophysical Research Letters 29, 2106 doi:10.1029/2002GL015633.

St. Laurent, L., Stringer, S., Garrett, C., Perrault-Joncas, D., 2003. The generation of internal tides at abrupt topography. Deep-Sea Research 50, 987–1003.

ELSEVIER

Deep-Sea Research II 51 (2004) 3043–3068

DEEP-SEA RESEARCH
PART II

www.elsevier.com/locate/dsr2

Internal wave generation in a global baroclinic tide model

Harper L. Simmons[a,b,*], Robert W. Hallberg[b], Brian K. Arbic[c]

[a]*International Arctic Research Center, University of Alaska Fairbanks, Fairbanks, AL 99775, USA*
[b]*Geophysical Fluid Dynamics Laboratory, P.O. Box 308, Princeton, NJ 08542-0308, USA*
[c]*Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ 08544-0710, USA*

Abstract

The energy flux out of barotropic tides and into internal waves ("conversion") is computed using a global domain multi-layer numerical model. The solution is highly baroclinic and reveals a global field of internal waves radiating way from generation sites of rough topography. A small number of sites where intense internal wave generation occurs accounts for most of the globally integrated work done on the barotropic tide and dominates sites such as the Mid-Atlantic ridge. The globally integrated conversion of the M_2 barotropic tide is 891 Gigawatts and the globally integrated rate of working of the ocean by astronomical forcing is 2.94 Terawatts. Both of these estimates are close to accepted values derived from independent methods. Regional estimates of conversion are also similar to previous inferences, lending additional confidence that the solution has captured the essential physics of low-mode internal wave generation and that numerical prediction of conversion has skill in regions where no previous estimates are available.
© 2004 Elsevier Ltd. All rights reserved.

1. Introduction

Tidal forcing by the sun and moon acts as a body force on the ocean, and it responds by accelerating as a barotropic flow. Internal wave generation subsequently occurs as a result of the flow of the stratified ocean over topographic variations such as ridges, trenches, seamounts, and shelf breaks (e.g. Bell, 1975; Baines, 1982; Hibiya, 1986). This generation of internal waves corresponds to a baroclinic wave drag and results in a transfer of energy out of the barotropic tide and into internal (baroclinic) waves. In this paper, we will refer to this flux as a "conversion". The area integrated conversion (Joules per second or Watts) is a rate of work, and has the same units as dissipation. The spatial redistribution of energy by propagating internal waves will be referred to as a baroclinic energy flux, and has units of Watts per meter. It should be understood that the energy extracted from the barotropic tide is not truly lost from the earth–sun–moon mechanical energy budget until it has been converted to heat by molecular viscosity. Some of the internal waves that are excited have high wavenumber and shear,

*Corresponding author. International Arctic Research Center, University of Alaska Fairbanks, Fairbanks, AL 99775, USA.
E-mail address: hsimmons@iarc.uaf.edu (H.L. Simmons).

0967-0645/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dsr2.2004.09.015

generate turbulence, and hence are dissipated by viscosity quickly. Others have lower wavenumber and are relatively stable (St. Laurent and Garrett, 2002). These low wavenumber waves are rapidly organized into a modal structure and are free to radiate away from the generation site, provided they are not excited at their critical latitude where their group velocity is zero. There is also more immediate viscous dissipation in turbulent bottom boundary layers. Observational evidence for the radiation of tidal frequency low-mode internal waves over thousands of kilometers comes from satellite altimetry (Ray and Mitchum, 1996, 1997; Ray and Cartwright, 2001; Cummins et al., 2001), current meters (Hendry, 1977; Chiswell, 2002; Alford, 2003) and acoustics (Dushaw et al., 1995). The cascade of energy from the barotropic tide though internal tides to turbulence is the central topic of the HOME experiment (Rudnick et al., 2003).

We are interested in the internal tides because they are a contributor to the internal wave continuum (Munk, 1981). The dissipation of tidal energy is also of inherent interest because of the implications for classical geophysical problems such as the length of the day and the evolution of the earth–moon system. Astronomical inference constrains the total dissipation of tidal energy in the earth–sun–moon system to be 3.7 Terawatts (TW), with 3.5 TW dissipated in the ocean (Munk, 1997; Egbert and Ray, 2001). Analysis of the tides using analytical methods, altimetry, and tidal modeling suggests that approximately 2.5 TW of this energy loss occurs through turbulent dissipation by direct dissipation in bottom boundary layers, primarily in shallow waters, and does not contribute to interior diapycnal mixing.

Observed turbulent diffusivities in the ocean interior range from $0.1\,cm^2\,s^{-1}$ in the thermocline (Ledwell et al., 1993, 1998) to greater than $10\,cm^2\,s^{-1}$ in the deep ocean above rough topography (Polzin et al., 1997; Ledwell et al., 2000). The main source of this energy is the internal wave field of the ocean. Munk and Wunsch (1998) calculated that 2 TW must be dissipated throughout the ocean to sustain a bulk turbulent diffusivity of $1cm^2\,s^{-1}$, maintaining the observed density stratification against an assumed overturning rate [although this amount has been disputed by Webb and Suginohara (2001)]. One Terawatt of energy loss from the barotropic tide has been attributed to internal tide generation, the subject of this paper. Previously, the global distribution of energy flux into the internal tides has been examined using theory (Sjöberg and Stigebrandt, 1992; Morozov, 1995), and tidal models (Kantha and Tierney, 1997; Egbert and Ray, 2000, 2001; Jayne and St. Laurent, 2001). Observational studies have shown that mixing rates are enhanced in regions of internal tide generation (Kunze and Toole, 1997; Lueck and Mudge, 1997; Polzin et al., 1997; Ledwell et al., 2000; Lien and Gregg, 2001; Pinkel et al., 2001; Kunze et al., 2002; Moum et al., 2002).

Previous estimates of the spatial distribution of energy sinks for the barotropic tides have used data-assimilative hydrodynamic models or theory. Morozov (1995) estimated 1.1 TW for the M_2 tide. Egbert and Ray (2001) estimated 0.70 ± 0.15 TW for the M_2 tide and 1 ± 0.25 for all lunar and solar constituents, but suggested that their estimate is conservative. Egbert and Ray (2001) have described the loss of energy from the barotropic tides as a energy sink term in the shallow water equations (see Section 2.3). Misfit between the hydrodynamic model and the assimilated altimeter data was attributed to "dissipation" of the barotropic tide by unresolved internal wave generation. This results in an estimate of total dissipation of the barotropic tide, but does not differentiate between the unresolved generation of low-mode internal waves of tidal frequency ("internal tides") and direct frictional dissipation. Furthermore, while Egbert and Ray's estimate of energy input is compelling and is strongly constrained by the altimeter data, it is not a *direct* calculation of the energy input into internal tides because those dynamics are precluded from their barotropic model, and their estimate hinges on the premise that their model misfit was primarily due to the missing physics of internal tide generation, and not due to intrinsic model error.

Baines (1973, 1982) estimated the conversion rate analytically. He considered a rotating, stratified, linearized, inviscid fluid, and decomposed the flow into a baroclinic component and a barotropic

component with equivalent sea surface height expression. This reduces the problem to the determination of the solution for the barotropic flow and the computation of the additional flow required to satisfy the original stratified dynamics and boundary conditions. The form of this additional term was interpreted as a "body force" on the barotropic tide. Baines calculated the adiabatic body forcing of the internal tides for simple solutions for the tidal flow in idealized two-dimensional geometries.

There have been a large number of regional numerical simulations of baroclinic tides. Studies include Kang et al. (2000), who studied baroclinic energy fluxes away from Hawaii using a regional two layer tide model, Kowalik and Polyakov (1999), who studied diurnal tides over Kashevarov Bank, as well as studies by Cummins and Oey (1997), Xing and Davies (1998), Holloway and Merrifield (1999), Cummins et al. (2001), and Merrifield et al. (2001). Niwa and Hibiya (2001) diagnosed barotropic to baroclinic conversion in the Pacific Ocean in an ocean model using the analytical Baines body force. Quantitative comparisons with results from many of these studies will be discussed in Section 4.

Although there have been a number of previous attempts to quantify the global input of energy into the internal wave field of the ocean, there have been no previous direct calculations of either the global internal wave field or the conversion rate that could unambiguously implicate internal wave generation by the tides as being responsible for the missing TW of energy loss from the barotropic tide. Our goal is to investigate the phenomena of barotropic-to-baroclinic mode conversion, internal wave generation, and energy redistribution by baroclinic energy flux in a process-oriented manner, using a high resolution baroclinic tide model. To this end, we will consider an initially quiescent, stratified ocean. Significantly, the ocean will be adiabatic. The motivation for this is that we wish to avoid any parameter dependence on sub-grid-scale parameterization of turbulence, we wish to consider the internal wave generation in the adiabatic limit, and we hypothesize that adiabatic low-mode internal wave generation can account for the TW of energy loss from the barotropic tide. While the model does have viscous stresses, it is adiabatic in the sense that there is no diapycnal mixing and hence the water mass census remains fixed for all time. In a companion study (Arbic et al., 2004), we also explore baroclinic tides in a global model. The distinction between that study and the present one is that in Arbic et al. (2004) the accuracy of the surface tides is the goal, and baroclinicity is shown to improve the fidelity of simulations of the surface tide, whereas here our motivation is to focus on the internal wave generation itself.

2. The vertically integrated primitive equations

Because the barotropic tides and their scattering of energy to internal waves is naturally represented by a modal description, an isopycnic coordinate ocean model is a natural tool for investigating the energetics of internal tides. We use the Hallberg Isopycnal Model (Hallberg and Rhines, 1996). Fig. 1 illustrates the relationship between the isopycnal coordinate system used by the model and the geopotential coordinate system. In isopycnal coordinates $z_{1/2} = \eta$ is the free surface displacement, and the subsequent interface elevations are measured from the resting sea surface height ($z = 0$). Interface thickness, h_n is the positive definite distance between adjacent interfaces, i.e., $h_n = z_{n-1/2} - z_{n+1/2}$. The ocean floor $-H(x, y)$ is rigid, though for our subsequent development of an energy equation this is not necessary. Tidal forces and loading do in fact

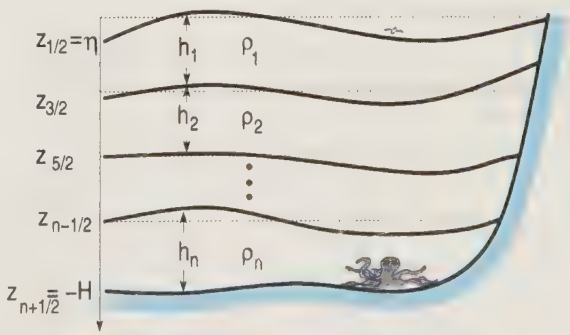


Fig. 1. Schematic for shallow water equations.

cause the ocean floor to deflect, and this should be accounted for in our energy equations because of the associated modification of the reference level for calculation of potential energy. Here, we account for the deformation of the ocean bottom in the traditional manner, via the Love numbers and a scalar approximation to self-attraction and loading (Section 2.1), and take the ocean bottom as fixed in time. This is appropriate at this stage for our process-oriented approach to the energetics of the tides.

The Boussinesq, adiabatic, hydrostatic, and nonlinear shallow water momentum equations are written for each layer as

$$\frac{\partial}{\partial t}\boldsymbol{u}_n + (\zeta_n + f)\hat{\boldsymbol{k}} \times \boldsymbol{u}_n = -\nabla\left(M_n + \frac{1}{2}\boldsymbol{u}_n^2 + g\Pi\right) + \boldsymbol{H}_n + \boldsymbol{V}_n. \quad (1)$$

$\zeta_n = \hat{\boldsymbol{k}} \cdot (\nabla \times \boldsymbol{u}_n)$ is the relative vorticity. M_n is the nth layer perturbation Montgomery potential, which is the layer integrated analogue to perturbation geopotential pressure. It is defined recursively as

$$M_1 = g_{1/2}\eta, \quad M_{n+1} = M_n + g_{n+1/2}\, z'_{n+1/2},$$

where $g_{n+1/2} = g(\rho_{n+1} - \rho_n)/\rho_n$, and $z'_{n+1/2}$ is the departure of the interface from its resting depth. For the first interface ($n = \frac{1}{2}$), $g = g_{1/2}$ is the full gravity. Tidal forcing, $g\Pi$, is in the form of the gradient of a potential. Details of this forcing are described in detail in Section 2.1. Momentum is diffused vertically by $\boldsymbol{V}_n$, a quadratic drag at the ocean bottom, and horizontally by $\boldsymbol{H}_n$, a biharmonic diffusion operator.

The continuity equation for each layer takes the from of a thickness tendency equation and is written as

$$\frac{\partial}{\partial t}h_n + \nabla \cdot (\boldsymbol{u}_n h_n) = 0. \quad (2)$$

Details of the finite difference numerical solution to (1) and (2) can be found in Hallberg and Rhines (1996) and Hallberg (1997). We will use the split time stepping scheme of Hallberg (1997). This scheme splits (1) into a computationally inexpensive but rapidly evolving barotropic mode, and a slowly evolving but computationally expensive baroclinic mode. These modes are defined in Section 2.3.

2.1. Tidal forcing

Tidal forcing (Π) arising from the eight primary constituents (M2, S2, N2, K2, K1, O1, P1, Q1) (Gill, 1982 and Table 1) has been added to the momentum equations. The formulation follows Marchuk and Kagan (1984), by considering a tide generating potential $g\Pi_{\mathrm{eq}}$ with corrections due to both the earth tide $(1 + k - l)$ and a scalar approximation to self-attraction and loading ($\alpha\eta$). Thus the full tidal potential in (1) is

$$g\Pi = g(1 + k - l)\Pi_{\mathrm{eq}} + g\alpha\eta. \quad (3)$$

The *equilibrium tide*, Π_{eq}, arises from the astronomically derived gravity producing forces. It is modified by several factors. The Love numbers, k and l account for the reduction of the ocean tide

Table 1
Frequencies (ω), Love numbers (k and h), and amplitude (a) functions for the eight principle constituents of tidal forcing

Constituent	Name	Origin	ω ($\times 10^{-4}$ s^{-1})	$1 + k - l$	a (m)
1	K_1	Luni-solar declinational	0.7292117	0.736	0.141565
2	O_1	Principal lunar declinational	0.6759774	0.695	0.100661
3	P_1	Principal solar declinational	0.7252295	0.706	0.046848
4	Q_1	Larger lunar elliptic	0.6495854	0.695	0.019273
5	M_2	Principal lunar	1.405189	0.693	0.242334
6	S_2	Principal solar	1.454441	0.693	0.112743
7	N_2	Larger lunar elliptic	1.378797	0.693	0.046397
8	K_2	Luni-solar declinational	1.458423	0.693	0.030684

because of the yielding of the solid earth by tidal forces (l) and from the perturbation of the potential due to the redistribution of mass within the earth caused by the yielding (k). The Love numbers are frequency dependent, with $1+k-l$ generally close to 0.7 (Wahr, 1981). The term α in (3) provides a scalar approximation to the modification of the ocean's tidal response as a result of self-attraction and loading (SAL) (Hendershott, 1972). Rigorous determination of SAL requires an extremely cumbersome integration over the earth surface, rendering a proper treatment of the tidal equations an integro-differential equation (Ray, 1998). Limitations of the scalar approximation to SAL are discussed by Ray (1998), who concluded that the scalar approximation introduces phase errors of up to 30° and amplitude errors of 10% into a global scale tidal simulation.

The equilibrium tide is written for the jth diurnal tidal constituent as

$$\Pi_{\mathrm{eq},j} = a_j \sin 2\phi \cos(\omega_j t + \lambda) \tag{4}$$

and for the jth semidiurnal constituent as

$$\Pi_{\mathrm{eq},j} = a_n \cos^2 \phi \cos(\omega_j t + 2\lambda), \tag{5}$$

where ϕ is latitude and λ is longitude. The frequencies (ω_j), amplitudes (a_j) and Love numbers are listed in Table 1. Recognizing that Eqs. (4) and (5) require the computationally expensive evaluation of trigonometric functions at every grid point and every time-step, tidal forcing is introduced in the following mathematically equivalent form. Making use of the identity $\cos(A+B) = \cos(A)\cos(B) - \sin(A)\sin(B)$, we can write the eight tidal forcing constituents as

$$\begin{aligned}\Pi_{\mathrm{eq}} = \sum_{j=1}^{4} [\beta_j \quad a_j \quad \cos^2\phi [\cos(\omega_j \quad t)\cos 2\lambda \\ - \sin(\omega_j \quad t)\sin 2\lambda] \\ + \beta_{j+4} a_{j+4} \sin 2\phi [\cos(\omega_{j+4} t)\cos 2\lambda \\ - \sin(\omega_{j+4} t)\sin 2\lambda]],\end{aligned} \tag{6}$$

which allows all the trigonometric functions of ϕ and λ to be pre-computed. Note that we have written $\beta_j = 1 + k_j - l_j$. When mode-splitting is employed, tidal forcing is added only to the barotropic equations. In Section 2.3, we will show analytically that tidal forcing should not appear in the baroclinic mode.

2.2. Energetics of the vertically integrated primitive equations

We wish to construct an energy budget in order to determine the work done by the tides and friction. For simplicity we will use the horizontally continuous equations here, but the constructed budget will actually be calculated from the horizontally discrete terms in the model. Our budget is constructed through the operation $h_n \boldsymbol{u}_n \cdot (1) + \frac{1}{2}\boldsymbol{u}_n^2 \cdot (2)$, and gives

$$\begin{aligned}\boldsymbol{u}_n h_n \frac{\partial \boldsymbol{u}_n}{\partial t} + \frac{1}{2}\boldsymbol{u}_n^2 \frac{\partial h_n}{\partial t} \\ = -\boldsymbol{u}_n h_n \nabla M_n - g\boldsymbol{u}_n h_n \nabla \Pi + \boldsymbol{u}_n h_n \\ \times \left[(\zeta_n + f)\hat{\boldsymbol{k}} \times \boldsymbol{u}_n - \nabla\left(\frac{1}{2}\boldsymbol{u}_n^2\right)\right] - \frac{1}{2}\boldsymbol{u}_n^2 \nabla \cdot (\boldsymbol{u}_n h_n) \\ + \boldsymbol{u}_n h_n \boldsymbol{H}_n + \boldsymbol{u} h_n V_n.\end{aligned} \tag{7}$$

Note that for our Boussinesq model, a constant factor of ρ_0 (the mean density of the ocean) will need to be multiplied through our energy density equation in order to obtain the traditional units for energy. Because the model conserves both mass and momentum, a budget constructed in this manner will close by construction. The two terms on the left-hand side are the time tendency of kinetic energy density. The first term on the right-hand side is the kinetic-to-potential energy conversion. The next term is the work done on the ocean by the tidal forces. The next three terms are the Coriolis and advection terms. The final two terms are the energy sinks due to the horizontal and vertical viscous stresses.

We can construct an available potential energy budget through the operation $M_n \cdot (2)$, giving us

$$M_n \frac{\partial}{\partial t} h_n + M_n \nabla \cdot (\boldsymbol{u}_n h_n) = 0. \tag{8}$$

Adding (7) and (8) gives a total energy equation.

At this point a number of manipulations of an energy budget based on vector and product rule identities appropriate to the continuous equations could be made such as recasting it into

flux-divergence form, but since our guiding principle is to construct an energy budget which applies strictly to our model momentum and continuity equations, these traditional manipulations are foregone. Note also that the Coriolis term is retained. While it drops out of the continuous equations point-wise because of a simple vector identity, this is not true of the discrete equations. The Coriolis terms here use the Sadourny (1975) energy conserving discretization, which would identically vanish upon global integration if a non-split time-stepping scheme were used. With a split time stepping scheme, the global cancellation is not perfect, but this term is six orders of magnitude smaller than the leading terms. The discrete energy budget has been constructed following the derivation of the energy budget from the continuous equations and it balances by construction. However, unless a trapezoidal time stepping scheme is used, the time-discretized first two terms in (7) are not guaranteed to exactly agree with the time derivative of thickness weighted kinetic energy density, i.e.,

$$\boldsymbol{u}h\frac{\partial \boldsymbol{u}}{\partial t}+\frac{1}{2}\boldsymbol{u}^2\frac{\partial h}{\partial t}\neq\frac{1}{2}\frac{\partial h\boldsymbol{u}^2}{\partial t}.$$

In a non-split case, these differences are small and result purely from temporal truncation errors. In the split-explicit case, there is an additional complicating factor. With a split time stepping scheme, there are two estimates of the free surface height—one from the sum of the layer thicknesses, the other from stepping forward the external mode equation. A positive definite continuity solver must be used for the layer thicknesses, but as there is no positive definite continuity solver that is linear in the layer thicknesses, the two estimates do not agree. The difference between these two estimates must be reconciled somehow—either by dilating all of the layers (Bleck and Smith, 1990), iterating the barotropic solution (Paul Schopf, pers. comm.), or by adding a fictitious mass source to the barotropic solution that sets up a divergent barotropic flow field an drives the two solutions toward each other, as is done here. This latter solution is efficient and has worked quite well for simulations of low-frequency flows. Unfortunately, the difference radiates as short-wavelength external gravity waves, and this shows up in the energetic analysis as a spurious energy source in the time tendency of kinetic energy. In hindsight, it would have probably been preferable to use the diabatic solution of dilating the layers to match the external calculation (Bleck and Smith, 1990). Fortunately this energy source is still only a few percent of the tidal forcing, and these waves tend to be of such short wavelengths that they are effectively dissipated by the horizontal viscosity. Further, as they have frequencies that are higher than those of the shortest resolved internal waves (and in our many-layer case higher than the buoyancy frequency), there is no linear conversion to baroclinic waves—the primary focus of this study. Based upon comparisons between split and un-split simulations we find that this fictitious energy source does not have a qualitative impact on the tidal solutions and analysis presented here, provided the degree of splitting is eight-to-one or less.

2.3. Modal energy equations

In the previous section, we transformed the model's momentum equations into a total kinetic energy budget in order to quantify the balance between energy input by tidal forces and dissipation by friction. In order to quantify the energy input into the baroclinic internal waves by inviscid, adiabatic processes, we now derive *modal* energy equations for an inviscid version of our layered system. For simplicity, we will work with the linearized momentum equations. It is straightforward to include the nonlinear terms (ζ_n and $\boldsymbol{u}^2$), but is more algebraically complicated. We neglect them a posteriori in our derivation as these terms are $O(u^3)$ in our energy equation. Furthermore, these terms were diagnosed to be two orders of magnitude smaller than the linear terms over most of the domain. We define the barotropic mode velocity and Montgomery potential as their thickness weighted average. For velocity, this is

$$\bar{\boldsymbol{u}}=\frac{1}{\Sigma_{n=1}^{k}h_n}\sum_{n=1}^{k}h_n\boldsymbol{u}_n, \tag{9}$$

where k is the total number of layers. The baroclinic mode is the departure from the thickness weighted average ($\boldsymbol{u}'_n = \boldsymbol{u}_n - \bar{\boldsymbol{u}}$). By this definition the thickness weighted vertical averaged of a baroclinic quantity vanishes.

Substituting $\boldsymbol{u}_n = \bar{\boldsymbol{u}} + \boldsymbol{u}'_n$ and $M_n = \overline{M} + M'_n$ into (1) and (2), neglecting $O(u^3)$ and viscous terms, we obtain a budget for the barotropic mechanical energy density for each layer through the operation $h_n\bar{\boldsymbol{u}} \cdot (1) + \overline{M} \cdot (2)$. Summing over all layers and eliminating any terms that will vanish upon vertical summation, this yields

$$\sum_{n=1}^{k}\left[\bar{\boldsymbol{u}}h_n\frac{\partial\bar{\boldsymbol{u}}}{\partial t} + \overline{M}\frac{\partial h_n}{\partial t}\right] = \sum_{n=1}^{k}[-g\bar{\boldsymbol{u}}h_n\nabla\Pi - \nabla\cdot(\bar{\boldsymbol{u}}h_n\overline{M}) - \bar{\boldsymbol{u}}h_n\nabla M'_n]. \quad (10)$$

Similarly, a baroclinic mechanical energy budget is obtained through the operation $h_n\boldsymbol{u}'_n \cdot (1) + M'_n \cdot (2)$, yielding

$$\sum_{n=1}^{k}\left[\boldsymbol{u}'_n h_n\frac{\partial\boldsymbol{u}'_n}{\partial t} + M'_n\frac{\partial h_n}{\partial t}\right] = \sum_{n=1}^{k}[-\nabla\cdot(\boldsymbol{u}'_n h_n M'_n) + \bar{\boldsymbol{u}}h_n\nabla M'_n]. \quad (11)$$

The term

$$\boldsymbol{J} = \sum_{n=1}^{k}\boldsymbol{u}'_n h_n M'_n \quad (12)$$

is the depth-integrated baroclinic energy flux vector. The term

$$C = \sum_{n=1}^{k}\bar{\boldsymbol{u}}h_n\nabla M'_n \quad (13)$$

is leading order depth-integrated barotropic-to-baroclinic energy conversion, and is balanced by the same term, but of opposite sign in the barotropic energy density equation (10). Note that the tidal forcing, $\boldsymbol{u}'_n h_n\nabla\Pi$, is eliminated from our depth-integrated baroclinic energy budget (11) by our definition of baroclinicity—the tidal forces do not input energy into the baroclinic mode.

3. Description of numerical experiments

The vertical coordinate is potential density referenced to 2000 m (σ_2). We explored using a realistic stratification as an initial condition, but found that even after a month the ocean was still experiencing large-scale adjustment that precluded a simple extraction of the tidal signal. Although we will continue to explore this approach, it proved to introduce too many challenges at this stage. For this reason, we elected to set the initial condition to a state of rest, with a single, horizontally uniform stratification profile (zero available potential energy), based on a zonal mean of the Subtropical North Pacific (Fig. 2) computed from the 1998 World Ocean Atlas (Antonov et al., 1998; Boyer et al., 1998). This stratification is representative of most of the world ocean, but clearly cannot represent high latitudes well.

From this profile, we configured two and 10 layer experiments. The two layer experiments were constructed by averaging σ_2 from our profile into two bins, one for waters from the surface to 1100 m in depth, and the other for waters from 1100 m to the ocean bottom, so that active interfaces lie at the surface and 1100 m depth.

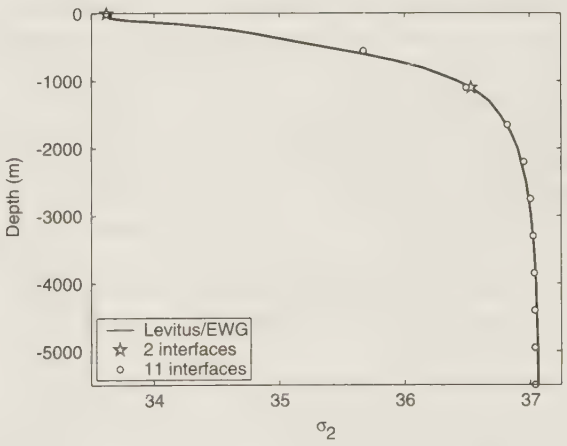


Fig. 2. Zonal mean potential density referenced to 2000 m (σ_2) for the Pacific, with two (star) and 10 (circles) layer densities and interface depths indicated. The two-layer solution has two active interfaces, one at the free surface, the other at 1100 m. The 10-layer solution has 11 active interfaces that are separated by 550 m layers.

For 10 layer experiments, the profile was constructed by equipartitioning σ_2 into 550 m thick layers. In this latter case, the third interface lies at the same depth, 1100 m, as the second interface in the two-layer experiments. By choosing equal thickness layers, we ensure that topography at all depths can interact with "local" stratification even though this would not be the optimal distribution for representing the flat-bottom normal modes, as might be suggested by Wentzel–Kramer–Brillouin (WKB) stretching. The rationale for considering two layers is that this system, by construction, cannot support higher modes and there is an interest in the relatively simple and energetically dominant gravest internal mode response. A two layer system is also computationally inexpensive. The 10 layer system is still computationally feasible for a single constituent and can be used to examine higher-mode conversions.

The horizontal grid covers the globe with a uniform spherical grid at one-eighth degree resolution from 78°S to 65°N. Above this latitude, the grid transitions to a bipolar grid (Murray, 1996) that allows polar regions to be resolved without a coordinate singularity occurring over the ocean. A characteristic of this grid is that the ocean resolution north of 65°N is nearly a uniform rectangular grid with a horizontal grid resolution close to the longitudinal resolution of the spherical grid at 65°N. Additional details of the model configuration can be found in Appendix A.1. Our simulations are all started from a state of rest. Runs presented in the main body of the text are forced with only a single semidiurnal constituent, M_2. A solution forced with a single frequency obviates the need to average over long period "beats" that occur (such as fortnightly tides as well as longer periods) when multiple constituents are included. The M_2-only configuration is of particular interest, as it is the dominant constituent by approximately a factor of two in amplitude (Table 1), and there is an extensive literature devoted to this constituent. Because the tides in the deep ocean are nearly linear, it is fortunately possible to learn much from running with a single constituent. Nonetheless, we acknowledge that interactions between components is an important physical process. Some discussion of experiments with eight constituents will occur in Appendix A.2, but more extensive discussion will be deferred to a future publication. While many more experiments were conducted in order to assess the sensitivity of C and $\boldsymbol{J}$ to various model parameters, forcing and resolution, our two and ten layer experiments capture much of the essential physics. Details of the additional configurations and their results can be found in Appendix A.2.

4. Results

A practical idealization necessary for this study was the use of a horizontally uniform stratification. While this eliminates transient wave adjustment, and allows us to examine the baroclinic wave generation in an otherwise resting ocean, it distorts the physics of conversion and internal wave energy flux. We can think of our horizontally uniform ocean as having been subjected to a WKB-like normalization such that the thermocline depth and non-rotating gravity wave speed are horizontally constant. WKB normalization is commonly used in the presentation of observational data as a convenience to facilitate comparison between regions of differing stratifications (Leaman and Sanford, 1975; Althaus et al., 2003). In order to approximately "correct" our estimates to account for the observed variation of topography, we will apply an a posteriori "renormalization" of energy flux and conversion by

$$(\hat{C}, \hat{\boldsymbol{J}}) = (C, \boldsymbol{J}) \left(\frac{\Delta\rho(x,y)}{\Delta\rho_0} \right)^{1/2} \tag{14}$$

and isopycnal displacement by

$$\hat{\eta} = \eta \left(\frac{\Delta\rho_0}{\Delta\rho(x,y)} \right)^{1/4}. \tag{15}$$

$\Delta\rho_0$ is the potential density difference between the upper and lower layer in our two-layer model and $\Delta\rho(x,y)$ is the climatological map of spatial varying potential density difference between waters above 1100 m and those below. The caret denotes renormalized quantities that might have been realized had the model been run with a

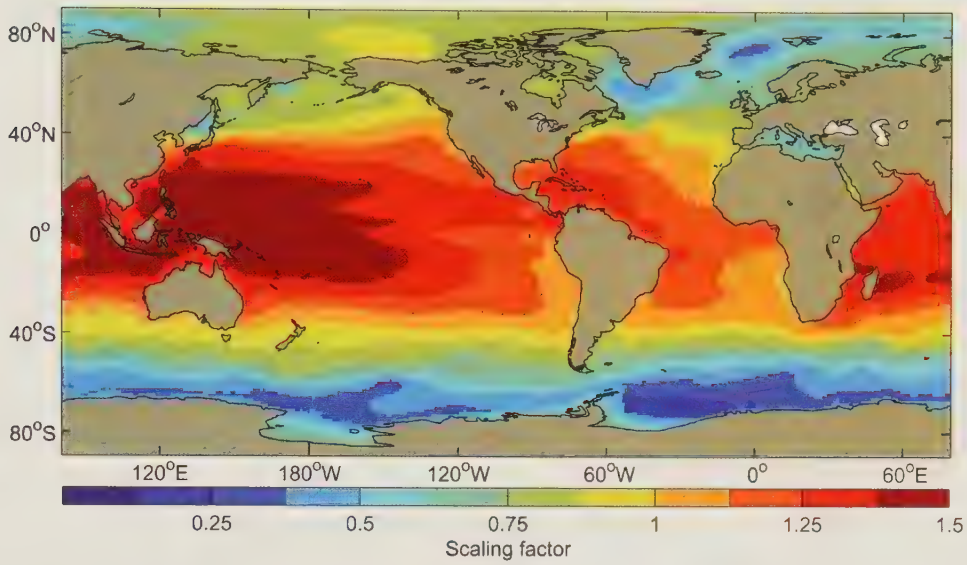


Fig. 3. Map of the multiplicative scaling factor used in (14) used to adjust our diagnosed energy flux and conversion rate for the neglected geographic variation in stratification. Note that this scaling factor is actually fairly modest, with the zonal mean value ranging from 0.4 to 1.3. To adjust internal isopycnal displacement we use the inverse square-root of this field (i.e., Eq. (15)).

non-uniform stratification. Subsequently reported conversion rates and fluxes will be rescaled using (14), and isopycnal displacements will be rescaled using (15). A map of the energy scaling factor (Eq. (14)) is shown in Fig. 3. Note that the magnitude of the renormalization is actually fairly modest with the zonal mean value of the scaling factor in (14) ranging from 0.4 to 1.3.

4.1. Generation of internal waves

The integration begins from a state of rest. We find that the tides energetically equilibrate within about 15 days, or about 30 M_2 periods (Fig. 4). The model is integrated for 21 days, with our analysis based on an average over the last two tidal cycles. Horizontal friction is relatively weak and once the model has equilibrated, the average tidal work over several tidal cycles is approximately balanced by the vertical friction in the system (Table 2). In the absence of numerical mode-splitting (Section 2.3) the balance between energy input by tidal forcing (T) and dissipation (ΣF) was found to be nearly exact. With a split time stepping scheme, the imbalance is still less then 10%. The small mismatch between forcing and dissipation is balanced by the small but spurious time tendency of kinetic energy term (the first term in Eq. (7)) alluded to in Section 2.3. When all terms in (7) are included, the residual imbalance is six orders of magnitude less than the leading order steady-state balance between forcing and dissipation.

The main features of the barotropic tides are set up within a few tidal periods, with basin scale Kelvin waves propagating around the ocean basins (Fig. 5). As the ocean sloshes back and forth across rough topography, internal waves are generated (Figs. 6–8), which propagate away from numerous hot-spots. These waves are clearly seen as an isopycnal displacement, with amplitudes of order 20 m, but in some selected regions the displacements are an order of magnitude larger. After 20 days, a complicated "sea" of internal waves has saturated the ocean basins. One of the key features of the fully developed internal wave field are the large wavelengths approaching the critical latitudes ($\omega = f$ at 74.5° for the M_2 tide). This will be discussed further in Section 4.3. Note the complicated basin scale interference patterns,

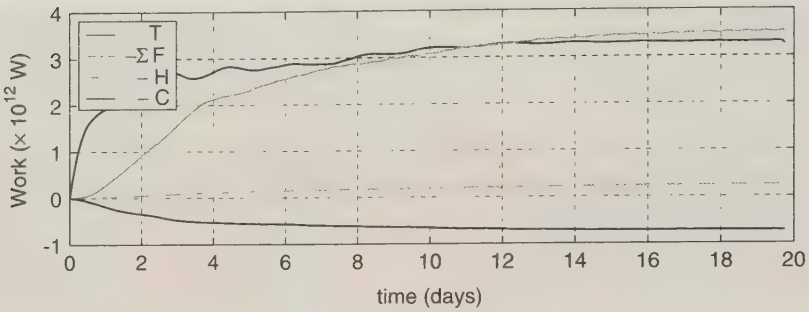


Fig. 4. Time-series of various terms in global energy budget. The data has been low-pass filtered to remove the (considerable) semidiurnal variability. The labels in the legend are T, the total work by the tidal forcing, ΣF, the total work done by friction, $-H$, the work done by horizontal friction, and C, the barotropic to baroclinic conversion rate. Dissipation is multiplied by negative one here for the purpose of visualization, so that if the traces of T and $-\Sigma F$ were exactly coincident, they would be in perfect balance.

Table 2
Globally integrated renormalized convergence rate ($\hat{C}$), work by vertical (V) and horizontal (H) friction, and work done by tides (T)

EXP	$\hat{C}$ (GW)	V (TW)	H (TW)	T (TW)
Two layer	671	3.18	0.20	3.17
Ten layer	891	2.71	0.45	2.94

Units for $\hat{C}$ are Gigawatts, units for V, H, and T are Terawatts.

and the tidal beams that appear to maintain their signatures over thousands of kilometers.

4.2. Barotropic tides

The barotropic tides are the forcing for our internal tides and therefore it is appropriate to assess their fidelity. A simple measure of the quality of the barotropic tidal solution can be made via a comparison between the computed surface tides and results from a very accurate data-assimilative model, TPX0.6 (Egbert and Erofeeva, 2002). TPX0.6 has absolute root-mean-square (RMS) discrepancies for the M_2 tide of less than 3 cm in water deeper than 1000 m. Fig. 5 shows M_2 tidal maps from the two-layer solution and from TPX0.6. Although certain features of the tidal solution are different, the overall amplitude and phase is qualitatively similar. For example, the major amphidromic systems of the Pacific, Indian, and Atlantic oceans are in the right locations, and the spatial pattern of tidal amplitudes are reproduced well. The main region of discrepancy is the Southern Ocean. The wiggles in the phase that are particularly evident in our solution in the North Pacific are the surface manifestation of internal waves. They are also seen in altimeter data (see, e.g., Dushaw, 2002 and references therein), but are not in TPX0.6, which assimilates a low-pass filtered version of the altimeter data, and uses a barotropic model for hydrodynamic interpolation.

Our two-layer barotropic solution are more energetic than observed, with correlations of $R^2 = 0.91$ for amplitude and $R^2 = 0.60$ for phase. Our solution is of comparable accuracy to that of forward models of tides without a gravity wave drag scheme, but is less accurate than forward models with a gravity wave drag parameterization that has been carefully tuned (Jayne and St. Laurent, 2001; Arbic et al., 2004; Egbert et al., 2004). We believe that for the goals of this study (i.e. a process-oriented investigation of internal tide generation) the quality of our solutions are sufficient.

4.3. Baroclinic response

The strong baroclinicity of the tidal response is clearly evident in a section of meridional velocity shown in Fig. 9. The velocity in the upper ocean is in many places 180° out of phase with the barotropic current, and it is surface intensified. Near the ocean bottom, in the vicinity of

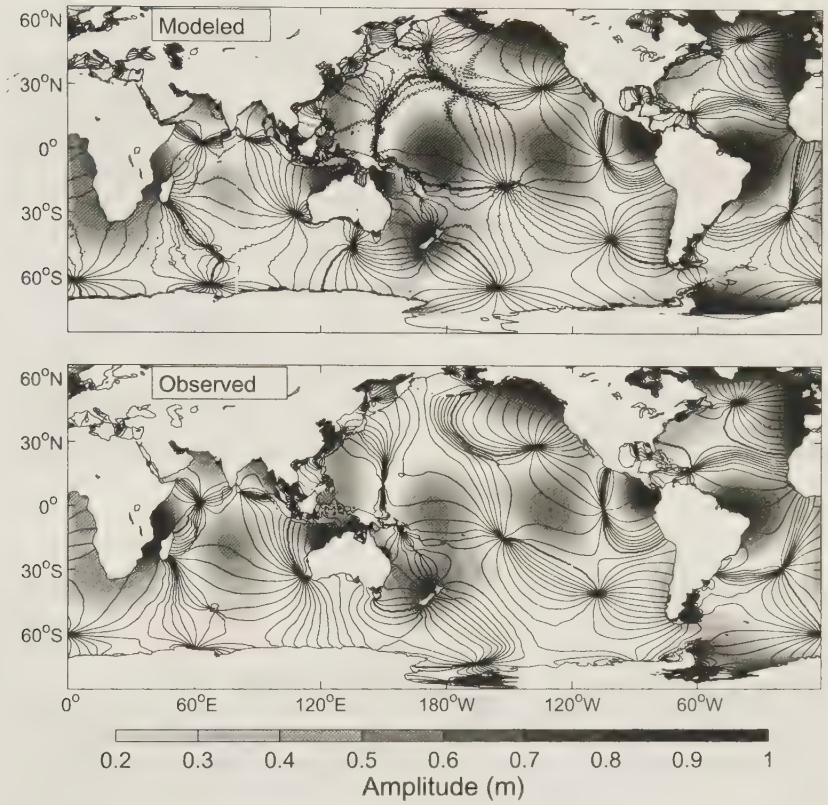


Fig. 5. M_2 tides from our M_2 only solution ("Modeled"), and from TPX0.6 ("Observed"), a very accurate data-assimilative tide model.

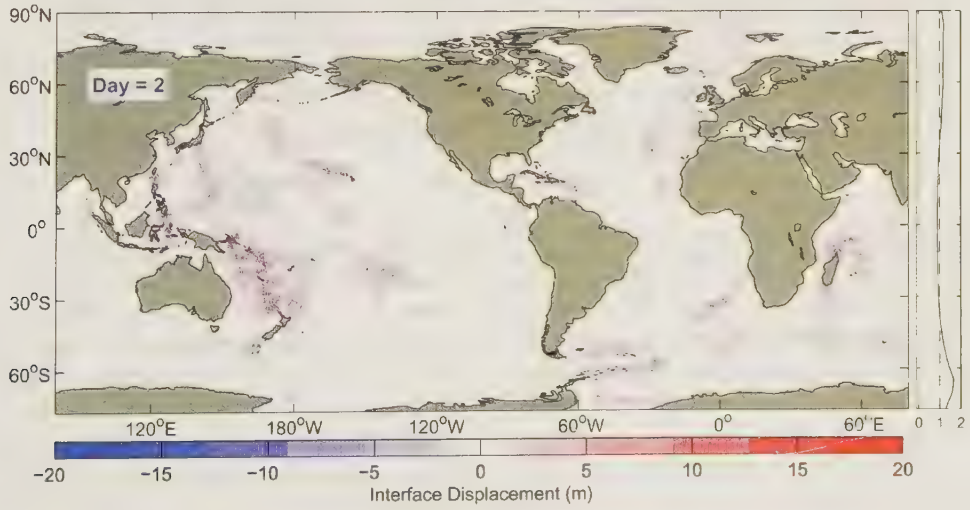


Fig. 6. Interface displacement, $\hat{\eta}$, normalized according to Eq. (15), on day two of spinup of the two-layer M_2 simulation. The resting depth of the interface is at 1100 m. The zonal mean of the normalization factor is shown on the right side of the plot.

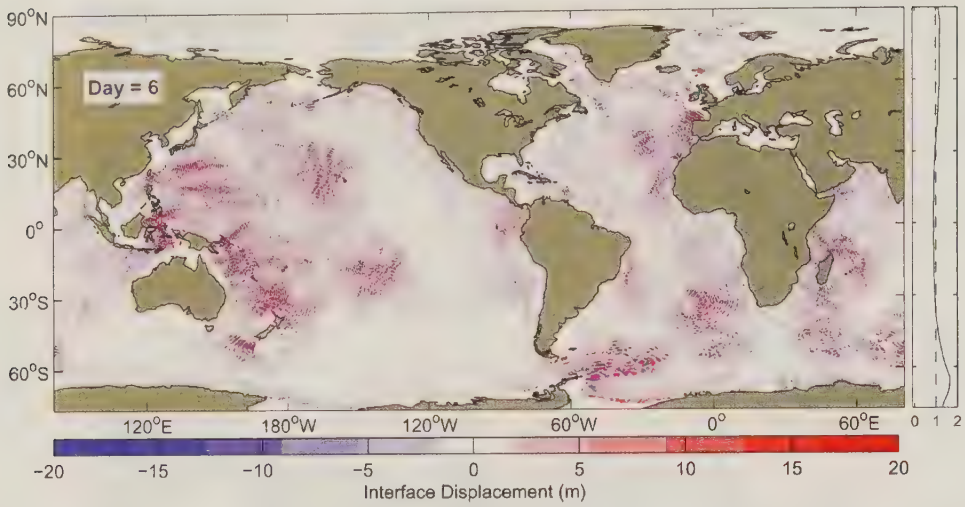


Fig. 7. Interface displacement, $\hat{\eta}$, normalized according to Eq. (15), on day six of spinup of the two-layer M_2 simulation. The resting depth of the interface is at 1100 m. The zonal mean of the normalization factor is shown on the right side of the plot.

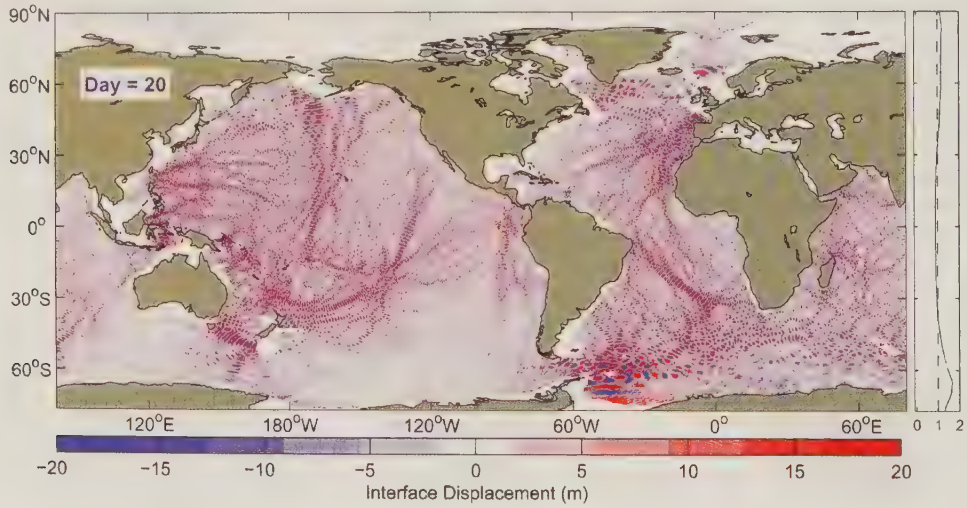


Fig. 8. Interface displacement, $\hat{\eta}$, normalized according to Eq. (15), on day 20 of spinup of the two-layer M_2 simulation. The resting depth of the interface is at 1100 m. The zonal mean of the normalization factor is shown on the right side of the plot.

roughness elements, small-scale disturbances to isopycnals are evidence for a low energy, higher mode response. A plan view from our two-layer experiment, comparing the depth-integrated flow with the upper layer flow is shown in Fig. 10. The striking difference between the upper layer flow (right panel) and the barotropic flow (left panel) is another indication of the baroclinicity of the tidal response. Note that while the barotropic flow is evident in the surface flow, it is mostly obscured by the many regions where the flow direction alternates direction at the wavelength of the internal tides, with larger magnitude than the large-scale barotropic flow. In these regions a comparison between a shallow water equations tidal solution and in situ current meter data could be as much as

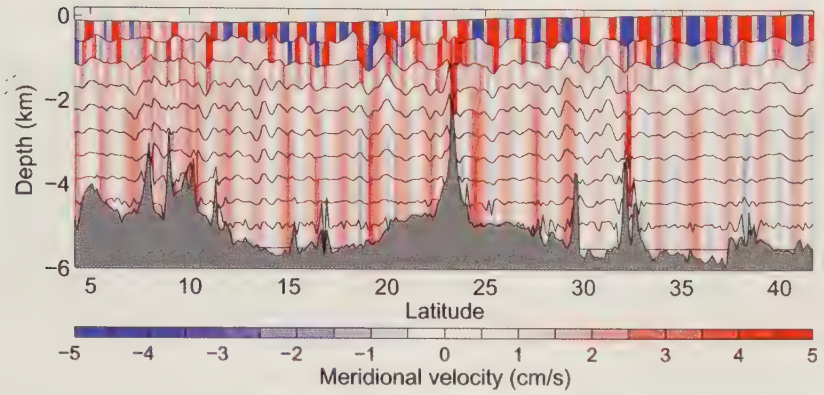


Fig. 9. Meridional section of meridional velocity in the Hawaiian Islands through Kauai Channel (approximately 159°W) for our 10 layer solution. Velocity is shaded between bounding isopycnals. Relative to the vertical scale of the plot, the ocean surface has been given a vertical scale exaggeration of 500, and interior isopycnals are exaggerated by a factor of 10.

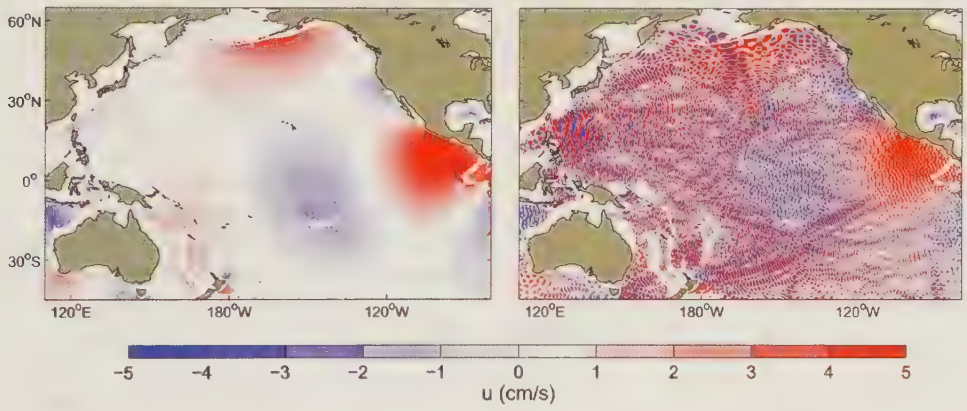


Fig. 10. Zonal component of velocity at day 20 of the two-layer M_2 simulation. Left panel: depth-integrated flow. Right panel: upper layer flow. For the purpose of visualization velocity has been set to zero for regions where water depth is less than 2200 m, so that we are only viewing flow where the lower layer is at least as thick as the upper layer (1100 m).

180°out of phase, and hence agreement would be very poor. Kowalik and Polyakov (1999) made such a comparison in the Sea of Okhotsk, and found reasonable agreement for current *amplitudes*, but found that direction was completely out of *phase* when compared with a shallow-water equations tide solutions. Agreement with observations was excellent when baroclinicity was included.

Another interesting feature of the solution is the small spatial scales evident in the interfacial waves near the equator (see, e.g., Fig. 8), and the larger wavelengths at higher latitudes. Note for example the large spatial scales in the vicinity of Drake Passage and in and around the Labrador Sea. The dispersion relationship for internal-mode gravity waves in the rotating, linear, hydrostatic two-layer system is given by

$$||k|| = \sqrt{\frac{(\omega^2 - f^2)H}{g' h_1 h_2}}, \tag{16}$$

where k is the wavenumber, H is the total ocean depth, g' is the reduced gravity of the two-layer system, and ω is the frequency of the wave of interest. The analytical M_2 internal wavelength ranges from about 70 km at the equator to about 300 km at about 10° from its critical latitude, where $\omega = f$. The dominant wavelength in our M_2-only two-layer solution, crudely estimated from the maximum of a smoothed spectral estimate of the zonal wavenumber at 10° intervals, is plotted over the theoretical curve (Fig. 11). The close agreement suggests that the latitudinal variation in wavenumber is well explained by the linear dispersion relation. Thus our computational grid resolves the semidiurnal baroclinic waves with approximately 10 grid-points at the equator. The K_1 analytical wavelength has also been plotted to illustrate that diurnal frequencies are resolved even better. The ability to resolve all waves increases with latitude. Note the increasing wavelength approaching the critical latitude in both the analytical curves and the modeled internal wave field. Very similar results are found in our 10-layer solution (not shown) indicating that the response is dominated by the first baroclinic mode, even with 10 layers. Space precludes a deeper investigation into the modal response, though it is clearly warranted. Also, while the requisite physics are present in our solution, we have not investigated physically important processes that can evolve our internal wave spectrum such as parametric subharmonic instability (Olbers, 1983; McKinnon and Winters, 2004).

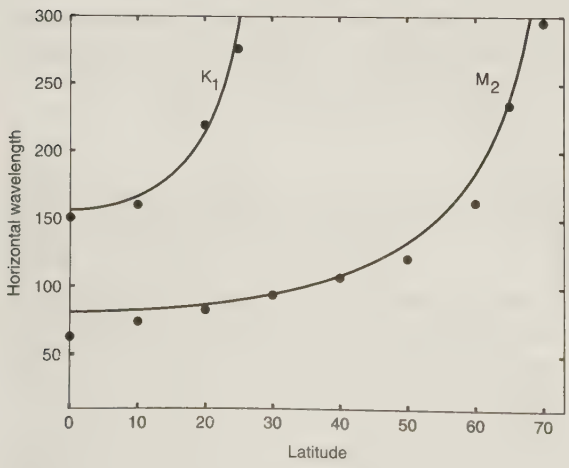


Fig. 11. Theoretical wavelength vs. latitude (solid lines) and wavelength of internal waves (circles) from two layer solution.

4.4. Global conversion and tidal rate of work

The renormalized global conversion rate, $\hat{C}$ is 0.67 TW (Table 2) for our two-layer M_2 experiment. Increasing the number of layers to 10 raises this estimate 0.89 TW, an increase of 33%. The work by the M_2 tide on the ocean for our two layer cases is 3.17 TW and is 2.94 TW for our 10-layer case (less work being consistent with a higher conversion rate with 10 layers), with the 10 layer case closer to accepted values. These values are somewhat larger than the 2.54 TW reported by Egbert and Ray (2001), but the agreement is encouraging considering that we are using a tide model that does not assimilate observed tides.

4.5. Regional estimates

Conversion rates from selected regions are labelled in Fig. 12 and tabulated in Table 3. We would like to stress that the place names we have assigned regions are generally for convenience, and are not necessarily meant to imply that a specific geological feature is responsible for the conversion, though in some cases this is certainly the case. For instance "Azores" really indicate a region of rough topography of which the Azores are only a component. In Fig. 13 we present a global map of the renormalized baroclinic energy flux vectors, $\hat{\boldsymbol{J}}$ overlaying a map of barotropic to baroclinic conversion rate. The relationship between internal wave generation, internal wave propagation (see Figs. 6–8), and baroclinic energy flux divergence is clear. Note that the fluxes in the figure have been smoothed with a Hanning window of 2×2 degrees and then sub-sampled in order to give a meaningful large-scale picture of the general pattern and magnitude of fluxes. This smoothing eliminates the many narrow and sinuous beam patterns evident in Fig. 8. Later, we will show a regional map of baroclinic energy flux where the rich details in the flux field are more evident.

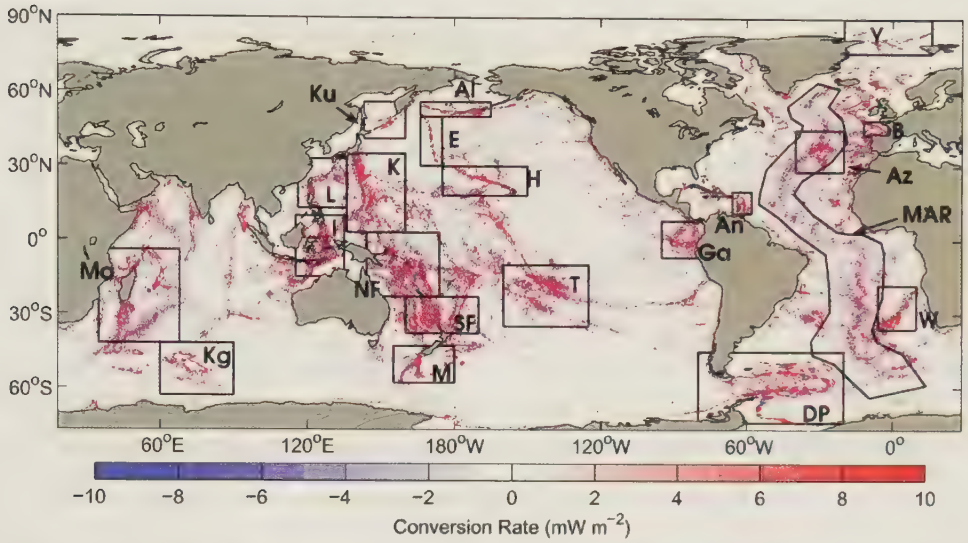


Fig. 12. M_2 conversion rate $\hat{C}$ overlayed with regions where area integrated conversion is calculated. Values are tabulated in Table 3. Key: Al = Aleutians Islands/Trench, An = Antilles, Az = Azores, B = Bay of Biscay, DP = Drake Passage/Scotia Sea, E = Emperor Seamounts, SF = South Fiji basin, Ga = Galapagos, H = Hawaii, I = Indonesia, K = Kyushu/Palu Ridge, KG = Kerguelan Plateau, Ku = Kuril Islands/Trench, M = Macquarie Ridge, MA = Madagascar, MAR = Mid-Atlantic Ridge, L = Luzon Strait, NF = North Fiji Basin, T = Tuomoto, W = Walvis Ridge, Y = Fram Strait/Yermak Plateau. Note that the MAR excludes the region delineated by the Azores.

Table 3
Conversion rates, in gigawatts, for various regions of interest (shown in Fig. 12), with comparisons to previous estimates, where available

EXP	Al	An	Az	B	DP	E	Ga	H	I	KG	Ku	K	L	M	MA	NF	SF	T	W	Y	∑	EPR	ABB	MAR	GLOBAL
Two layer (raw)	15	5	19	15	79	3	13	17	54	10	4	25	9	35	25	42	51	40	17	1	478	0.06	0.14	14	648
Ten layer (raw)	19	7	20	14	99	3	14	22	56	15	6	29	16	36	45	55	48	50	19	2	576	0.11	1.46	42	867
Two layer (normalized)	12	7	19	12	34	3	15	23	76	6	3	35	13	28	32	60	55	53	18	1	504	0.07	0.16	12	671
Ten layer (normalized)	15	9	20	11	44	3	17	30	78	9	5	41	22	29	56	79	53	66	20	1	608	0.14	1.62	36	891
Niwa and Hibiya (2001)	10	—	—	—	—	—	—	15	85	—	—	21	41	—	—	39	—	23	—	—	—	—	—	—	
Merrifield et al. (2001)	—	—	—	—	—	—	—	9	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	
Kang et al. (2000)	—	—	—	—	—	—	—	5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	
Ray and Cartwright (2001)	—	—	—	—	—	—	—	6	—	—	—	—	—	—	—	—	—	11	—	—	—	—	—	—	—
Egbert and Ray (2001)	—	—	—	—	—	—	15	18	—	—	—	47	—	—	—	—	—	31	—	—	—	—	—	103	700 ± 150
Morozov (1995)	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1100
Jayne and St. Laurent (2001)	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1300
Sjöberg and Stigebrandt (1992)	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1070
St. Laurent and Garrett (2002)	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0.2	4	—	—
St. Laurent et al. (2003)	—	—	—	—	—	—	—	22	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—

The first two rows are the raw conversion rates from the model solution. The second two rows are the conversion rates renormalized by Eq. (14). Key: Al = Aleutians Islands/Trench, An = Antilles, Az = Azores, B = Bay of Biscay, DP = Drake Passage/Scotia Sea, E = Emperor Seamounts, Ga = Galapagos, H = Hawaii, I = Indonesia, K = Kyushu/Palu Ridge, KG = Kerguelan Plateau, Ku = Kuril Islands/Trench, M = Macquarie Ridge, MA = Madagascar, MAR = Mid-Atlantic Ridge, L = Luzon Strait, NF = North Fiji Basin, SF = South Fiji basin, T = Tuomoto, W = Walvis Ridge, Y = Fram Strait/Yermak Plateau, EPS = East Pacific Rise, GLOBAL = Global Ocean. The column marked ∑ is the sum of the previous columns. The EPR and ABB data are in units of mW m^{-2} and are in locations discussed in St. Laurent and Garrett (2002). Data for Egbert and Ray (2001) is taken from their Table 2, solution TPX0.4a.

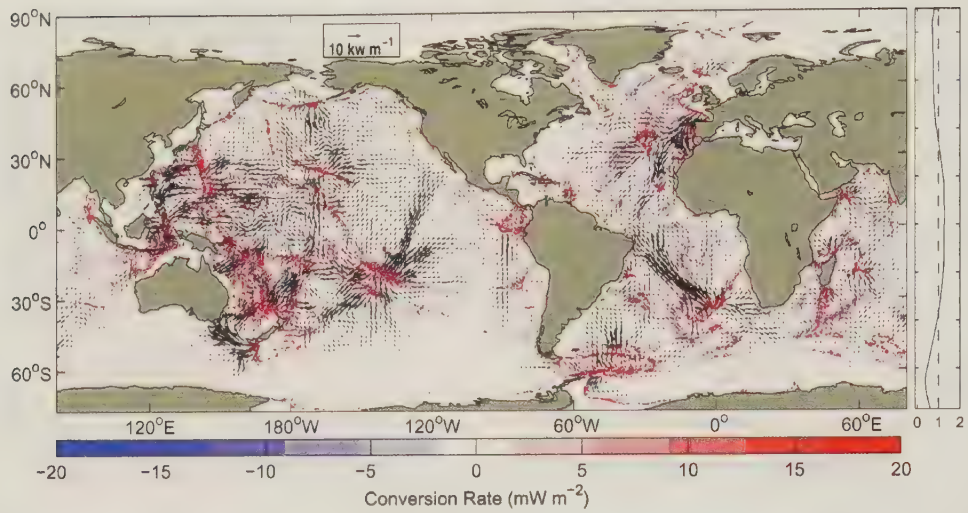


Fig. 13. Baroclinic energy flux $\hat{J}$ and conversion rate $\hat{C}$ from our two-layer simulation, normalized according to (14) and averaged over one tidal cycle. Only a subset of the flux vectors are shown. The zonal mean of the normalization factor (Eq. (14) and Fig. 3) is shown on the right side of the plot. Each vector component has been spatially smoothed with a bidirectional Hanning window with a width set to 2 degrees and then sub-sampled in order to give a meaningful large-scale picture of the general pattern and magnitude of fluxes. For visual clarity it was necessary to make the color range representing the conversion rate be different from Fig. 12, to clip vectors longer than $20\,\mathrm{kW\,m^{-1}}$, and to delete vectors shorter than $1\,\mathrm{kW\,m^{-1}}$.

The conversion is mostly positive (energy transfer is from barotropic to baroclinic motions), and in all of our regional estimates the net conversion is positive. However, over small spatial scales there are regions that are negative. For our two-layer simulations, the total barotropic-to-baroclinic conversion is $+1.40\,\mathrm{TW}$ and the total baroclinic-to-barotropic conversion is $-0.73\,\mathrm{TW}$, adding up to our global conversion of $+0.67\,\mathrm{TW}$. Regions of negative conversion were also found by Egbert and Ray (2001). In our simulation, negative conversion rates arise where there is a convergence of baroclinic energy flux, such as might hypothetically occur between two wavemakers that are an integer number of wavelengths apart and in phase so that the waves constructively interfere. One of the most energetic regions with respect to conversion is the Western Pacific (also noted by Niwa and Hibiya, 2001). The baroclinic energy fluxes emanating primarily northward from Tuomoto in the Pacific and from the Walvis Ridge in the South Atlantic are noteworthy, as is the large flux out of the Bay of Biscay. The marked asymmetry of the fluxes emanating from the Walvis Ridge is particularly interesting.

We note that a number of the regions that are significant generators of baroclinic waves have been commented on previously in the literature, and a global picture with many similarities to ours can be found in Egbert and Ray (2001), hereafter referred to as ER. The level of spatial detail of our conversion calculation allows us to now quantify fluxes for many different geographic features for which there are no previous estimates. The sparseness of the previous estimates in Table 3 reflects this increase in information. There are many regions and features that we could identify as internal wavemakers (our computational grid allows us to make estimates at literally millions of locations). In the interest of brevity, we have limited the number of regions for which we have made explicit estimates. Since we have explicitly simulated internal wave generation by the tides, this lends confidence to the contention of ER that their model misfit is due to the missing physics of baroclinic wave drag in their solution. We also note that the M_2 dissipation presented by ER (their plate 2 and their Table 2) also includes dissipation by quadratic bottom drag. Our map shows *only* the energy flux from the barotropic

mode going in to baroclinic motions (Eq. (11)). We find significant generation of internal waves in many of the regions considered "shallow" by ER (e.g., our Madagascar, Antilles, Luzon Strait, Macquarie Ridge, Kuril Islands, Kerguelan Plateau). For this reason it is difficult for us to compare our two maps in detail. Thus, we will compare with ER only in their "deep" regions where their drag is predominately attributed to baroclinic wave generation. In these regions, we find that our estimate is generally similar to ER's (our Fig. 12 and Table 3). For instance, our 10 layer solution produces a 30 GW conversion rate at Hawaii, as compared to 18 GW found by ER. In general, our estimates are somewhat larger than ER's, and our global, 10 layer M_2 conversion rate is 891 GW, compared to ER's 700 ± 150 GW. Note again that an absolute comparison of our two estimates is not possible because of their combination of work by bottom friction and by implied baroclinic wave drag. Nonetheless, we can speculate that the larger prediction in our solution may be partially attributable to a more energetic tide than ER (and than observed) and partially to ER's exclusion of shallow seas from their estimate. ER do suggest that their estimates may be conservative for this reason.

Our results are also quite similar to the Pacific basin computation of internal tide generation by Niwa and Hibiya (2001), with quantitative comparison of certain regions listed in Table 3. Another notable feature of our solution that is new to our knowledge, is the relatively large internal tide generation by the ridges and trenches surrounding the Scotia Sea. Here, we find as much as $36\,\mathrm{kW^{-1}\,m}$ of baroclinic wave energy radiating into the Atlantic, when averaged over an M_2 tidal cycle. The area-integrated conversion rate for the Drake Passage region is as much as 44 GW, or 5% of the global ocean. Naviera-Garabato et al. (2004) have recently inferred strong internal wave activity and mixing in the Scotia Sea, and have speculated as to the geostrophic vs. tidal origin of the observed internal energy in this region. While this study cannot speak for the geostrophic contribution, it does suggest that the internal tide contribution is non-trivial.

Another notable feature of our solutions is the strong conversion at the Walvis Ridge (18–20 GW) in the South Atlantic, with baroclinic energy fluxes radiating north-westward into the South Atlantic, and eastward towards the Indian Ocean. It seems that this site is included in ER's mid-Atlantic Ridge (MAR) estimate. A few of the many additional conversion sites of interest are the Bay of Biscay (15–17 GW), and Cape Farewell (South of Greenland in the Labrador Sea). Note also the generation at the Galapagos and the Lesser Antilles. Ray and Cartwright (2001) estimated that 11 GW are extracted from the barotropic tide at the Tuomoto Archipelago in the South Pacific, about a factor of four smaller than we (53–66 GW), ER (31 GW), or Niwa and Hibiya (23 GW) have found. Note that their estimate of 6 GW for Hawaii is similarly low. Ray and Cartwright suggest that their low estimate for Hawaii is due to poor sampling by the along-track altimeter data used in their study.

A detail of our baroclinic energy flux vectors for our two-layer solution superimposed over Hawaiian topography is shown in Fig. 14. The flux is confined to sinuous, narrow (several hundred kilometer wide) beams emanating irregularly along the ridge. These beams are a ubiquitous feature of our solution (Fig. 8). Chiswell (2002) estimated a baroclinic tidal energy flux of $4.4\,\mathrm{kW\,m^{-1}}$ at station Aloha (22.75 N, 158 W), north of Hawaii (indicated in Fig. 14). The flux at this location varies from 0.25 to $7\,\mathrm{kW^{-1}}$ to the north–west over a tidal cycle with a mean value of $3.5\,\mathrm{kW\,m^{-1}}$. The maximum flux along the Hawaiian Ridge is to the north–west, originating from French Frigate Shoals at approximately (24 N, 167 W), ranges over a tidal cycle from 10 to $45\,\mathrm{kW\,m^{-1}}$ with a mean value of $26\,\mathrm{kW\,m^{-1}}$. Rudnick et al. (2003) reported semidiurnal band fluxes at French Frigate Shoals of $8–34\,\mathrm{kW\,m^{-1}}$, which is very similar to what we have obtained here. We find Kauai Channel fluxes of $5–26\,\mathrm{kW\,m^{-1}}$ with a mean value of $17\,\mathrm{kW\,m^{-1}}$, where Rudnick et al. (2003) measured fluxes of $5–21\,\mathrm{kW\,m^{-1}}$. Kang et al. (2000) estimated 5.4 GW along entire Hawaiian ridge ($2.7\,\mathrm{kW\,m^{-1}}$) for M_2. Merrifield et al. (2001) found an integrated energy flux away from Hawaii of 9 GW using a baroclinic tide model, but the

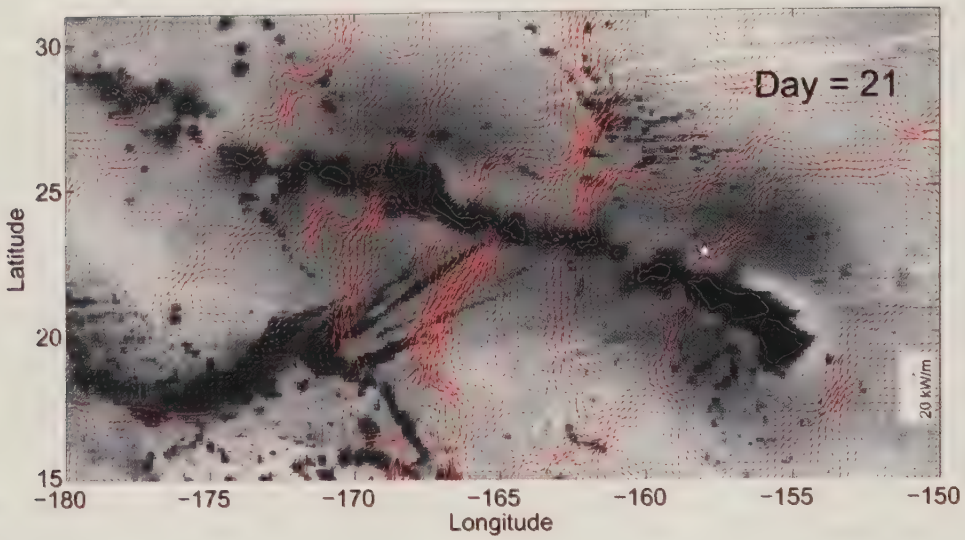


Fig. 14. Detail of baroclinic energy flux vectors near Hawaii for our two-layer solution. The star indicates the location of Station Aloha where Chiswell (2002) estimated a baroclinic tidal energy flux of 4.4 kW m ^{1}. Each vector component has been spatially smoothed with a bidirectional Hanning window with a width set to 0.25° and then every other vector is plotted. Underlying topography is the model topography at full resolution. A flux of 20 kW m^{-1} is indicated in the lower right-hand corner.

actual conversion was likely larger because they did not account for energy losses within their control volume. Their average baroclinic energy flux was 1.7 kW m^{-1} along the ridge, with a peak value of 10 kW m^{-1}. Dushaw (2002) estimated 2.6 GW for the Hawaiian ridge using altimetry, though he suggests that poor sampling makes this an underestimate. Using a linear theory, St. Laurent et al. (2003) have calculated internal wave generation for a number of abrupt geometries such as a knife-edge, a step, a top-hat ridge, and a top-hat trench. When they apply their theory to the Hawaiian Ridge, they find that as much as 11 kW m^{-1} of internal tide production, or 22 GW for an 2000 km ridge.

Another interesting feature of our solutions are the relatively modest contribution from the mid-ocean ridges. St. Laurent and Garrett (2002) used linear theory appropriate to the MAR, and found approximately 4 mW m^{-2} of conversion out of the barotropic M_2 tide at a site on the crest of the MAR. In this latter study, they find that mode-1 internal waves absorb less than 25% of the modal conversion, and that 30 modes are required to account for 80% of the conversion. Note, however, that the net predicted conversion at sites like this on the MAR is quite small because these high mode internal waves cannot add up to a significant conversion. Our 10-layer solution produces a modest amount of conversion at the same site as St. Laurent and Garrett (2002), with an energy flux of 1.5 mW m^{-2}. This is somewhat smaller than the 4 mW m^{-2} predicted by St. Laurent and Garrett (2002) for the total contribution from all modes, but our mode-1 dominated solution agrees well with their estimate that mode-1 carries 25% of the total flux. For their smooth East Pacific rise site we obtain 0.15 mW m^{-2} vs. St. Laurent and Garrett's value of 0.2 mW m^{-2}. Our East Pacific rise and Abyssal Brazil Basin estimates are given in Table 3, columns "ABB" and "EPR".

The MAR (excluding the Azores) is responsible for the conversion of 12 GW in our two-layer simulations, and 36 GW in our 10-layer solution, the latter accounting for approximately 3% of the global conversion in our 10-layer case. We find that the net conversion integrated over the MAR is comparable to generation sites like Hawaii. It is the huge area of ocean floor occupied by MAR topography that makes the contribution by the MAR at all significant to the global total. In Table 3, column "Σ" we have summed the

contribution for the 20 regions indicated in Fig. 12, but "Σ" does not include the MAR. The total for just these 20 regions is 75% of the global total conversion, yet these regions constitute less than 10% of the ocean sea floor area. Thus, we have found that the globally integrated conversion is dominated by internal wave generation at abrupt topography. In the supercritical theory (i.e., when the slope of internal wave characteristics exceed the slope of the topography) of St. Laurent et al. (2003), the spectrum is red, with 75% of the energy contained in mode-1. In the linear (subcritical) theory of St. Laurent and Garrett (2002) (a modification of Bell, 1975), the mode spectrum is proportional to the product of the bottom stratification, the integrated wavenumber spectrum (the bathymetric variance) and the tidal current. This theory is applicable to topography such as the MAR, predicting that the conversion peaks for mode-5. The conversion associated with this mode is very weak compared to sites where the peaks occurs at or near mode-1, e.g., Hawaii. The topographic roughness along the MAR is dominated by small scales, so that the absolute variance for topographic wavenumber capable of exciting a modal response is small compared to the variance associated with abrupt topographic features such as Hawaii. Thus, as in the linear theory, net conversion is roughly proportional to total topographic variance (Balmforth et al., 2002) and implies that regions like Hawaii are much more efficient at extracting energy from the barotropic tides. For regions like the MAR, mode numbers that are high with respect to our computational grid are required to capture a significant portion of the conversion, and we expect that our estimates will be low. It is interesting to note that our 10-layer solution, which can admit some higher modal response, does result in an 80% increase of conversion integrated over the MAR; in regions of abrupt topography, where generation is biased towards low modes, our solutions are generally very similar for both two and 10 layers. Presumably regions like the MAR will become more important as resolution increases, both vertically and horizontally, allowing a more realistic mode spectrum to be excited.

5. Summary and conclusions

A process-oriented study of the rate of conversion of the barotropic tide into baroclinic waves and the rate of working of the ocean by tidal forces has been conducted using a isopycnic coordinate ocean model. The accuracy of our tidal elevations is comparable to the accuracies found in un-tuned tidal simulations that do not assimilate data (Jayne and St. Laurent, 2001; Arbic et al., 2004; Egbert et al., 2004). With an overly energetic barotropic tide our tidal solutions produce a rich global field of internal waves. The baroclinic energy flux associated with propagating internal waves is concentrated in beams that extend thousands of kilometers from generations sites. We find a total of 2.94 TW of work is done by the M_2 tides in our 10-layer experiment, close to previous estimates. The globally integrated conversion with two layers is 0.67 TW. Our 10 layer estimate is 0.89 TW, which is close to the 0.7 ± 0.15 TW estimate of Egbert and Ray (2001), but note that they did not include many shallow regions (e.g., Fig. 13 and Table 3 denoted Madagascar, Lesser Antilles, Luzon Strait, Macquarie Ridge, Kuril Islands, Kerguelan Plateau) where we find significant internal wave generation.

Our solution captures many of the internal wave generation sites previously noted in the literature, and the model suggests many generation sites that to our knowledge have not been previously noted. Several such sites are the Scotia Ridge and Sandwich Islands and the Walvis Ridge in the South Atlantic. We are also able to make quantitative estimates of regional conversion that not have been made previously. Approximately 75% of the global conversion of barotropic tides into internal waves is accounted for by 20 internal wave generation sites (Table 3). In these hot-spots, the conversion rates and baroclinic energy fluxes compare well with previously available regional estimates. These are regions of abrupt topography, and dominate generation from sites such as the MAR. Thus conversion is dominated by low mode internal wave generation at regions of large net topographic variance such as trenches and island arcs. The fact that estimates compare favorably in regions where other estimates have been made give

us confidence that our solution has value as a prognostic tool for estimating fluxes in new regions. The quantitative estimate of the conversion is limited by the dependence on the local value of stratification. To account for spatially varying stratification, we have employed an a posteriori WKB-like renormalization (Eqs. (14) and (15)). Note that the magnitude of this renormalization is relatively modest and does not change the qualitative features of mode conversion, baroclinic wave generation and baroclinic energy flux.

The baroclinic response is primarily in the first internal mode even for our 10-layer solution. Evidence for this is the close agreement (for both the two and 10-layer solutions) between the internal wavelengths and those predicted by the two layer dispersion relationship. While we have not yet undertaken a full modal analysis of our 10-layer solution, there is some evidence that higher modes are being excited. Indications are the 33% higher conversion rates for our 10-layer solution (Table 3), the disturbance of isopycnals near roughness elements near the ocean bottom (Fig. 9), and the marked increase of conversion in the Abyssal Brazil Basin and East Pacific Rise. We regard the low-mode internal wave generation, and associated conversion from to barotropic to baroclinic motions to be represented in a physically robust manner. Our resolution, both horizontal and vertical, is suitable, by construction, to resolving the low modes. The increase of conversion at the Abyssal Brazil Basin and East Pacific Rise with increased vertical resolution are suggestions that as we increase horizontal and vertical resolution, these sites will become more important contributors. However, since we already obtaining nearly one TW of conversion with two layers, it is unlikely that the unresolved higher modes will account for a great deal more net conversion.

We have found that the adiabatic generation of *low-mode* internal tides is insensitive to many of the parameterizations that finite-difference models must employ (see our appendix), but are reluctant at this stage to make conclusions about the fraction of internal wave energy lost to local dissipation to that radiated far away from the generation site. We acknowledge that there are many aspects of the solution to which an ocean general circulation model are not well suited. The accurate determination of how much of energy extracted from the barotropic tides is dissipated locally versus the amount radiated away cannot be captured accurately by our adiabatic model because it cannot support the higher mode internal wave generation and turbulence that mediate this partitioning. With respect to this latter point, however, some theory suggests that low-mode waves are stable over times long relative to their fundamental frequency, and therefore our qualitative picture of adiabatic internal wave generation and energy redistribution by baroclinic energy fluxes radiating away from generation regions is likely a reasonable first-order description, even if the ultimate decay of these waves cannot be resolved explicitly for the foreseeable future.

Acknowledgements

HLS is indebted to Igor Polyakov for many helpful discussions of tide modelling, and to Louis St. Laurent for his encouragement and many discussions about internal tide generation over the last several years. Suggestions by the anonymous reviewers and by Eric Kunze and Chris Garrett are all gratefully acknowledged. Infrastructure support from IARC/Frontier and the Geophysical Fluid Dynamics Laboratory (GFDL) are acknowledged. Computations were carried out using GFDL's High Performance Computing facility, and essential code development and early experiments were conducted at the Arctic Regional Supercomputer Center. This work is supported in part by the National Science Foundation under Agreement No. OPP-0002239.

Appendix A

A.1. Details of model configurations

The basic experiments ("two-layer", "ten-layer") discussed in the main part of this paper had the following configuration. Horizontal resolution was one-eighth degree. Biharmonic isopycnal viscosity was set by a parameter, A, which is

multiplied by third power of the grid resolution. This parameter has units of velocity and was set to $0.01\,m\,s^{-1}$. The scalar self-attraction and loading parameter (SAL) was 0.094. The time step for the baroclinic mode was 80 s, and 10 s for the barotropic mode, i.e., a split timestep of eight-to-one. The non-dimensional quadratic bottom drag coefficient was 0.003. Vertical viscosity was set to the molecular value of $10^{-6}\,m^2\,s^{-1}$. Diapycnal mixing was zero in order to focus on the *adiabatic* generation of baroclinic internal waves.

We also conducted a number of additional simulations to assess the sensitivity of conversion rate to the following: horizontal resolutions of half, quarter, eighth, and sixteenth degree; specified constant diapycnal mixing ranging from 0 to $10^{-4}\,m^2\,s^{-1}$; biharmonic viscosity variations of two orders of magnitude, from 0.01 to $1.0\,m\,s^{-1}$; vertical viscosity ranging from $10^{-6}\,m^2\,s^{-1}$ (molecular) to $10^{-4}\,m^2\,s^{-1}$; filtering to reduce topographic roughness; bottom drag coefficient ranging from 0.003 to 0.3; SAL parameter of 0.084 or 0.094; either one (M_2 only) or the eight dominant tidal constituents. The barotropic time step for the four resolutions explored was 50, 25, 10, and 5 s for the half, quarter, eighth, and sixteenth degree numerical experiments, respectively. To assess the impact of split time stepping, the ratio of baroclinic to barotropic mode timesteps was additionally varied from sixteen-to-one to one-to-one.

Topography data was taken from the thirtieth-degree resolution ETOPO2 dataset (National Geophysical Data Center), down-sampled to the target grids ($\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}$ degree) by a simple nearest neighbor scheme without any pre-filtering, except when sensitivity to topographic roughness was investigated.

A.2. Sensitivity of results

Of the factors listed above, the only factors responsible for producing more than a 10% difference in global conversion rate are increasing horizontal and/or vertical resolution, filtering topography to reduce roughness, increasing the number of tidal constituents included in the forcing, and increasing the bottom drag coefficient. A bottom drag coefficient of 0.03, a factor of 10 time larger than is physically justifiable, had approximately a 10% effect on the globally integrated convergence. A coefficient of 0.3 reduced the conversion rate by a factor of two, and as such is significant. However, we feel that this is an unphysical parameter choice. Note, however, that it does result in more realistic barotropic tidal energy levels. This subject is addressed carefully by Arbic et al. (2004). Adjusting horizontal and vertical viscosity, diapycnal mixing and SAL all resulted in negligible changes to the solution. Based on these findings about parameter sensitivity, we will report only the effects of topographic roughness, vertical and horizontal resolution, and the number of tidal constituents.

A.2.1. Two layers

Global conversion more than doubles going from half to quarter degree resolution, but increases by less than 10% going to eighth degree. Going to one-sixteenth degree resolution only adds another 1%, so that our highest explicitly resolved conversion is 0.68 TW. The implication is that the topography that excites the mode-1 response has been adequately resolved at one-eighth degree. Higher horizontal resolution is required to resolve topographic wavenumbers capable of exciting a higher mode response, but, by construction, only mode-1 can be excited in a two layer system. Therefore, the 1% increase in convergence going from eighth to sixteenth degree resolution leads us to believe that the solution has converged in the physical (albeit not the mathematical) sense. Because of this apparent convergence, also we conclude that higher modes potentially excited by increasingly resolved high wavenumber topography are not spuriously projected into the gravest internal mode. We have extrapolated the two-layer case to zero grid spacing (the continuous limit). This gives us the estimate that the system removes 0.69 TW of energy from the barotropic M_2 tide.

A.2.2. Ten layers

The conversion increased by 30% going from quarter to eighth degree resolution, whereas our two-layer solution had less than a 10% increase

over this horizontal resolution transition. There is an additional 7% increase going to one-sixteenth degree. Our highest resolution (one-sixteenth degree horizontally, 10 layers vertically) explicitly resolved conversion is 1.00 TW. In the continuous limit, this conversion is extrapolated to 1.04 TW. The fact that the conversion rate increases with horizontal resolution when there are more than two layers is an indication that additional degrees of freedom are admitted with more layers. This higher mode number response can be seen in the section view of Fig. 9. The moderately increasing conversion (as compared to two-layer solutions) with respect to increasing horizontal resolution for these experiments suggests that these higher mode number fluxes are not entirely resolved.

A.2.3. Sensitivity to topographic roughness

Fig. 15 compares the wavenumber spectrum of filtered and unfiltered topography, used to investigate the sensitivity to bottom roughness. The filtering was a five-point bi-directional Hanning window applied both meridionally and zonally. For our basic experiments, direct interpolation of our source topography to our model grid by a nearest neighbor scheme most nearly preserves the roughness present in the source data, as measured by the topographic wavenumber spectrum (Fig. 15). Smoothing the source topography prior to regridding, as is common in terrain-following and geopotential coordinated models results in a loss of topographic variance at high wavenumbers and reduces the conversion rate by approximately 50% at one-quarter degree for both the two and 10-layer cases. However, at one-eighth degree the conversion rate is only reduced by 24%. This is evidence that the length scales responsible for generating mode-1 internal waves are better resolved by eighth degree resolution, so that filtering has less of an effect. At a quarter degree, the conversion rate reduction after filtering is the same for two and 10 layer solutions. This further suggests that the filtering at lower resolution has eliminated spatial scales responsible for generating higher mode response.

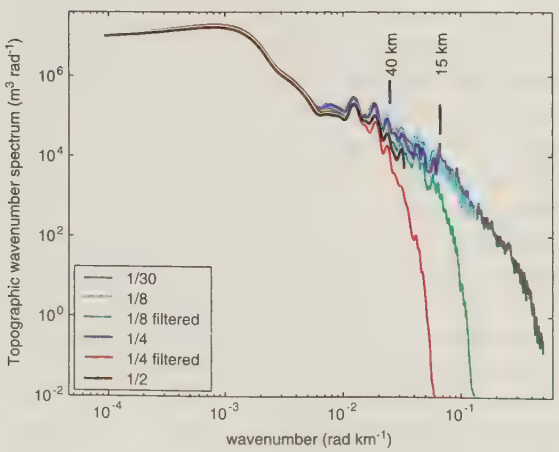


Fig. 15. Effect of spatial pre-filtering on wavenumber spectra of topography for a zonal section through Drake Passage. Note that the *spectrum* has been smoothed to improve clarity. The topography is downsampled through a simple nearest neighbor scheme.

A.2.4. Sensitivity of the work done by the tides

The work done by the tidal forcing, T, reduces with increased horizontal resolution, and is estimated to be 2.5 TW in the horizontally continuous 10-layer limit for the M_2 tide. The addition of eight constituents raises the global work by approximately 40%. Again assuming that the eight-constituent system converges like the M_2-only system, we estimate T for the eight primary tidal constituents to be 3.7 TW. Ten layers makes less than a 4% difference to the estimate. Inclusion of eight constituents raises T to 3.94 TW, which is approximately 10% larger than the accepted value of 3.5 TW.

A.2.5. Multiple constituent runs

With eight constituents, we find that the tides do not energetically equilibrate in the same manner as the single constituent case, and there is considerable variability in the globally integrated energy budget (Fig. 16) with clear spring-neap and longer period variations occurring. These longer period variations arise because of frequency beating of the forcing. The model is integrated for 60 days, with our analysis based on an average over the last three fortnightly cycles. It is interesting that the work done by frictional dissipation lags the work done by tides by several days. This time lag is likely to be sensitive to our parameterization of

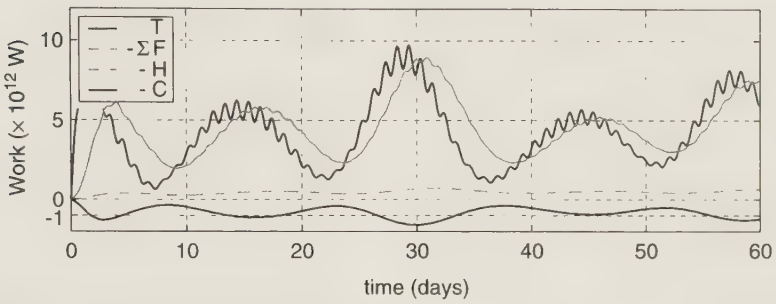


Fig. 16. Time-series of various terms in global energy budget for an eight constituent experiment. The data has been low-pass filtered to remove the semidiurnal variability. The labels in the legend are T, the total work by the tidal forcing, ΣF, the total work done by friction, H, the work done by horizontal friction, and C, the barotropic to baroclinic conversion rate.

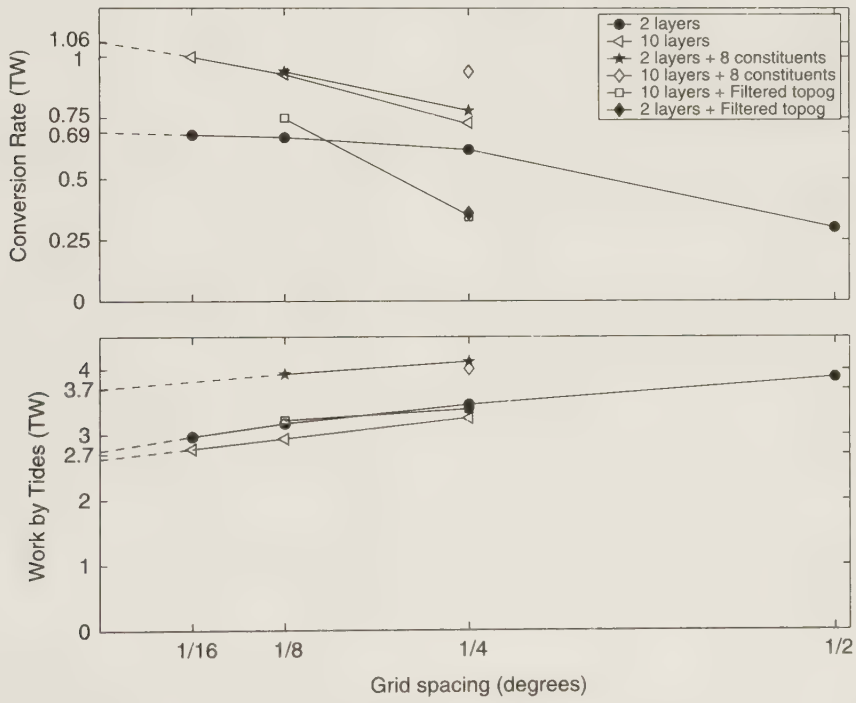


Fig. 17. Top panel: Conversion rate vs. horizontal resolution. Bottom panel: work by tide vs. horizontal resolution.

friction. The globally integrated conversion (discussed in Section 4.4) is also shown because it is interesting to note its long period modulation. The two-layer conversion is 0.94 TW, a 40% increase over our M_2-only solution. We can thus estimate that a 10-layer ocean would have a conversion of 1.3 TW. Assuming that the eight-constituent conversion has similar convergence properties with respect to horizontal resolution to that of M_2-only solutions, we estimate that in the two-layer horizontally continuous limit, the net conversion is 0.96 TW. Our estimate for the conversion in a 10 layer system is 1.34 TW, which is similar to, but slightly larger than the Egbert and Ray (2001)

value of 1 ± 0.25 TW estimate. Using the ratio of two-layer M_2-only to two-layer eight constituent tidal rate of work, we can estimate that the eight constituent 10 layer tidal rate of work to be 3.64 TW. Finally, a notable feature of the eight constituent experiments is a markedly increased conversion in the Western Pacific, particularly at Luzon Strait, where our conversion jumps from less than 22 GW for our 10-layer M_2 solution, to 46 GW for eight constituents.

A.2.6. Summary of sensitivity experiments

One of our goals of this study has been to establish some basic data about the resolution and computational expense necessary to meaningfully capture internal tide generation in a numerical model at the global scale. We have demonstrated the feasibility of global baroclinic simulations by showing that the conversion rates of barotropic to low-mode baroclinic tidal energy converge as a function of resolution. The first baroclinic mode appears to be adequately resolved at one-eight degree, and even one-quarter degree grids can produce meaningful results. Higher modes admitted by our 10-layer solution appear to be only partially resolved at one-sixteenth degree. Substantiated by the convergence in the conversion rate, the insensitivity to parameterization of viscous stresses and diapycnal mixing, and favorable comparison to independent estimates of conversion and baroclinic energy flux, it is suggested that numerical models can be a useful tool in understanding the internal tide generation problem (Fig. 17).

References

Alford, M.H., 2003. Redistribution of energy available for ocean mixing by long-range propagation of internal waves. Nature 423, 159–162.

Althaus, A., Kunze, E., Sanford, T., 2003. Internal tide radiation from Mendecino Escarpment. Journal of Physical Oceanography 33, 1510–1527.

Antonov, J.I., Levitus, S., Boyer, T.P., Conkright, M.E., O'Brien, T.D., Stephens, C., 1998. World Ocean Database 1998. vol. 4: Salinity of the Atlantic Ocean. Technical Report, NOAA Atlas NESDIS 30, US Government Printing Office, Washington, DC, 166 pp.

Arbic, B., Garner, S., Hallberg, R., Simmons, H., 2004. The accuracy of surface elevation in forward global barotropic and baroclinic tide models. Deep-Sea Research II, this issue [doi:10.1016/j.dsr2.2004.09.014].

Baines, P., 1973. The generation of internal tides by flat-bump topography. Deep-Sea Research 20, 179–205.

Baines, P.G., 1982. On internal tide generation models. Deep-Sea Research 29, 307–338.

Balmforth, N., Ierley, G., Young, W., 2002. Tidal conversion by subcritical topography. Journal of Physical Oceanography 32, 2900–2914.

Bell, T.H., 1975. Lee waves in stratified flows with simple harmonic time dependence. Journal of Fluid Mechanics 67, 705–722.

Bleck, R., Smith, L., 1990. A wind-driven isopycnic coordinate model of the North and Equatorial Atlantic Ocean. 1. model development and supporting experiments. Journal of Geophysical Research 95, 3273–3285.

Boyer, T.P., Levitus, S., Antonov, J.I., Conkright, M.E., O'Brien, T.D., Stephens, C., 1998. World Ocean Database 1998. vol. 1: Temperature of the Atlantic Ocean Technical Report, NOAA Atlas NESDIS 27, US Government Printing Office, Washington, DC, 166 pp.

Chiswell, S.M., 2002. Energy levels, phase, and amplitude modulation of the baroclinic tide off Hawaii. Journal of Physical Oceanography 32, 2640–2651.

Cummins, P., Oey, L., 1997. Simulation of barotropic and baroclinic tides off Northern British Columbia. Journal of Physical Oceanography 27, 762–781.

Cummins, P.F., Cherniawski, Y., Foreman, M., 2001. North Pacific internal tides from the Aleutian Ridge, altimeter observations and modeling. Journal of Marine Research 59, 167–191.

Dushaw, B., 2002. Mapping low-mode internal tides near hawaii using TOPEX/POSEIDON altimeter data. Geophysical Research Letters 29, 10.1029/2001GL013944.

Dushaw, B., Worcester, P., Cornuelle, B., Howe, B., 1995. Barotropic and baroclinic tides in the central North Pacific Ocean determined from long-range reciprocal acoustic transmissions. Journal of Physical Oceanography 25, 631–647.

Egbert, G.D., Erofeeva, S.Y., 2002. Efficient inverse modeling of barotropic ocean tides. Journal of Atmospheric and Oceanic Technology 19, 183–204.

Egbert, G., Ray, R., 2000. Significant dissipation of tidal energy in the deep ocean inferred from satellite altimeter data. Nature 405, 775–778.

Egbert, G.D., Ray, R.D., 2001. Estimates of M_2 tidal energy dissipation from TOPEX/POSEIDON altimeter data. Journal of Geophysical Research 106, 22475–22502.

Egbert, G.D., Ray, R.D., Bills, B.G., 2004. Numerical modelling of the global semidiurnal tide in the present day and in the Last Glacial Maximum. Journal of Geophysical Research 109, doi:10.1029/2003JC001973.

Gill, A.E., 1982. Atmosphere-Ocean Dynamics. International Geophysics Series, vol. 30. Academic Press, New York, 662 pp.

Hallberg, R., 1997. Stable split time stepping schemes for large-scale ocean modeling. Journal of Computational Physics 135, 54–65.

Hallberg, R., Rhines, P., 1996. Buoyancy-driven circulation in an ocean basin with isopycnals intersecting the sloping boundary. Journal of Physical Oceanography 26, 914–940.

Hendershott, M.C., 1972. The effects of solid earth deformation on global ocean tides. Geophysical Journal of the Royal Astronomical Society 29, 389–402.

Hendry, R.M., 1977. Observations of the semidiurnal internal tide in the western North Atlantic Ocean. Philosophical Transactions of the Royal Society of London A 286, 1–24.

Hibiya, T., 1986. Generation mechanism of internal waves by tidal flow over a sill. Journal of Geophysical Research 91, 7696–7708.

Holloway, P.E., Merrifield, M., 1999. Internal tide generation by seamounts, ridges, and islands. Journal of Geophysical Research 104, 25937–25951.

Jayne, S.R., St Laurent, L.C., 2001. Parameterizing tidal disspiation over rough topography. Geophysical Research Letters 28, 811–814.

Kang, S., Foreman, M., Crawford, W., Cherniawsky, J., 2000. Numerical modeling of internal tide generation along the Hawaiian ridge. Journal of Physical Oceanography 30, 1083–1098.

Kantha, L., Tierney, C., 1997. Global baroclinic tides. Progress in Oceanography 40, 163–178.

Kowalik, Z., Polyakov, I., 1999. Diurnal tides over Kashevarov Bank, Okhotsk Sea. Journal of Geophysical Research 104, 5361–5380.

Kunze, E., Toole, J.M., 1997. Tidally driven vorticity, diurnal shear, and turbulence atop Fieberling Seamount. Journal of Physical Oceanography 27, 2663–2693.

Kunze, E., Rosenfeld, L.K., Carter, G.S., Gregg, M.C., 2002. Internal waves in Monterey submarine canyon. Journal of Physical Oceanography 32, 890–1913.

Leaman, K., Sanford, T., 1975. Vertical energy propagation of inertial waves: a vector spectral analysis of velocity profiles. Journal of Geophysical Research 80, 1975–1978.

Ledwell, J.R., Watson, A.J., Law, C.S., 1993. Evidence for slow mixing across the pycnocline from an open-ocean tracer-release experiment. Nature 364, 701–703.

Ledwell, J.R., Watson, A.J., Law, C.S., 1998. Mixing of a tracer released in the pycnocline of a subtropical gyre. Journal of Geophysical Research 103, 21499–21529.

Ledwell, J.R., Montgomery, E.T., Polzin, K.L., Laurent, L.C.S., Schmitt, R.W., Toole, J.M., 2000. Mixing over rough topography in the Brazil Basin. Nature 403, 179–182.

Lien, R.-C., Gregg, M.C., 2001. Observations of turbulence in a tidal beam and across a coastal ridge. Journal of Geophysical Research 106, 4575–4591.

Lueck, R., Mudge, T., 1997. Topographically induced mixing around a shallow seamount. Science 276, 1831–1833.

Marchuk, G.I., Kagan, B.A., 1984. Ocean Tides: Mathematical Models and Numerical Experiments. Pergamon Press, Oxford.

McKinnon, J., Winters, K., 2004. Reconsidering a disqualified candidate for the energy balance of oceanic internal waves. EOS Transactions of the American Geophysical Union 84 (52), OS12G-07.

Merrifield, M.A., Holloway, P., Johnston, T., 2001. The generation of internal tides at the Hawaiian ridge. Geophysical Research Letters 28, 559–562.

Morozov, E., 1995. Semidiurnal internal wave global field. Deep-Sea Research I 42, 135–148.

Moum, J.N., Caldwell, D.R., Nash, J.D., Gunderson, G.D., 2002. Observations of boundary mixing over the continental slope. Journal of Physical Oceanography 32, 2113–2130.

Munk, W., 1981. Internal waves and small-scale processes. In: Warren, B.A., Wunsch, C. (Eds.), Evolution of Physical Oceanography—Scientific Surveys in Honor of Henry Stommel. MIT Press, Cambridge, MA, pp. 264–291.

Munk, W., 1997. Once again: once again-tidal friction. Progress in Oceanography 40, 7–35.

Munk, W.H., Wunsch, C., 1998. Abyssal recipes II: energetics of tidal and wind mixing. Deep-Sea Research I 45, 1977–2010.

Murray, R.J., 1996. Explicit generation of orthogonal grids for ocean models. Journal of Computational Physics 126, 251–273.

Naviera-Garabato, A.C., Polzin, K., King, B., Heywood, K.J., Visbeck, M., 2004. Widespread intense turbulent mixing in the Southern Ocean. Science 303, 210–213.

Niwa, Y., Hibiya, T., 2001. Numerical study of the spatial distribution of the M_2 internal tide in the Pacific Ocean. Journal of Geophysical Research 106, 22229–22441.

Olbers, D.J., 1983. Models of the oceanic internal wave field. Reviews of Geophysics 21, 1567–1606.

Pinkel, R., Munk, W., Worcester, P., et al., 2001. Ocean mixing studied near the Hawaiian Ridge. EOS Transactions of the American Geophysical Union 81, 545,553.

Polzin, K.L., Toole, J.M., Ledwell, J.R., Schmitt, R.W., 1997. Spatial variability of turbulent mixing in the abyssal ocean. Science 276, 93 -96.

Ray, R., Cartwright, D.E., 2001. Estimates of internal tide energy fluxes from Topex/Poseidon altimetry: Central North Pacific. Geophysical Research Letters 28, 1259–1262.

Ray, R., Mitchum, G., 1996. Surface manifestation of internal tides generated near Hawaii. Geophysical Research Letters 23, 2101–2104.

Ray, R., Mitchum, G., 1997. Surface manifestation of internal tides in the deep ocean: observations from altimetry and tide gauges. Progress in Oceanography 40, 135–162.

Ray, R.D., 1998. Ocean self-attraction and loading in numerical tidal models. Marine Geodesy 21, 181–192.

Rudnick, D., Boyd, T.J., Brainard, R.E., Carter, G.S., Egbert, G.D., Gregg, M.C., Holloway, P.E., Klymak, J.M., Kunze, E., Lee, C.M., Levine, M.D., Luther, D.S., Martin, J.P., Merrifield, M.A., Moum, J.N., Nash, J.D., Pinkel, R., Rainville, L., Sanford, T.B., 2003. From tides to mixing along the Hawaiian ridge. Science 301, 355–357.

Sadourny, R., 1975. The dynamics of finite-difference models of the shallow water equations. Journal of Atmospheric Science 32, 680–689.

Sjöberg, B., Stigebrandt, A., 1992. Computations of the geographical distribution of the energy flux to mixing processes via internal tides and the associated vertical circulation in the ocean. Deep-Sea Research 39, 269–291.

St. Laurent, L.C., Garrett, C., 2002. The role of internal tides in mixing the deep ocean. Journal of Physical Oceanography 32, 2882–2899.

St. Laurent, L.C., Stringer, S., Garrett, C., Perrault-Joncas, D., 2003. The generation of internal tides at abrupt topography. Deep-Sea Research I 50, 987–1003.

Wahr, J., 1981. Body tides on an elliptical rotating elastic and oceanless earth. Journal of the Royal Astronomical Society 64, 677–703.

Webb, D.J., Suginohara, N., 2001. The interior circulation of the ocean. in: Siedler, G., Church, J., Gould, J. (Eds.), Ocean Circulation and Climate, International Geophysics Series, vol. 77. Academic Press, New York, 715pp.

Xing, J., Davies, A.M., 1998. A three-dimensional model of internal tides on the Malin-Hebrides shelf and shelf edge. Journal of Geophysical Research 103, 27821–27847.

ELSEVIER

Deep-Sea Research II 51 (2004) 3069–3101

DEEP-SEA RESEARCH
PART II

www.elsevier.com/locate/dsr2

The accuracy of surface elevations in forward global barotropic and baroclinic tide models

Brian K. Arbic[a,*], Stephen T. Garner[b], Robert W. Hallberg[b], Harper L. Simmons[b,c]

[a]*Program in Atmospheric and Oceanic Sciences, Princeton University, P.O. Box CN710, Sayre Hall, Princeton, NJ 08544-0710, USA*
[b]*Geophysical Fluid Dynamics Laboratory/NOAA, Princeton, NJ 08542, USA*
[c]*International Arctic Research Center, University of Alaska Fairbanks, Fairbanks, AK, 99775, USA*

Abstract

This paper examines the accuracy of surface elevations in a forward global numerical model of 10 tidal constituents. Both one-layer and two-layer simulations are performed. As far as the authors are aware, the two-layer simulations and the simulations in a companion paper (Deep-Sea Research II, 51 (2004) 3043) represent the first published global numerical solutions for baroclinic tides. Self-consistent forward solutions for the global tide are achieved with a convergent iteration procedure for the self-attraction and loading term. Energies are too large, and elevation accuracies are poor, unless substantial abyssal drag is present. Reasonably accurate tidal elevations can be obtained with a spatially uniform bulk drag c_d or horizontal viscosity K_H, but only if these are inordinately large. More plausible schemes concentrate drag over rough topography. The topographic drag scheme used here is based on an exact analytical solution for arbitrary small-amplitude terrain, and supplemented by dimensional analysis to account for drag due to flow-splitting and low-level turbulence as well as that due to breaking of radiating waves. The scheme is augmented by a multiplicative factor tuned to minimize elevation discrepancies with respect to the TOPEX/POSEIDON (T/P)-constrained GOT99.2 model. The multiplicative factor may account for undersampled small spatial scales in bathymetric datasets. An optimally tuned multi-constituent one-layer simulation has an RMS elevation discrepancy of 9.54 cm with respect to GOT99.2, in waters deeper than 1000 m and over latitudes covered by T/P (66°N to 66°S). The surface elevation discrepancy decreases to 8.90 cm (92 percent of the height variance captured) in the optimally tuned two-layer solution. The improvement in accuracy is not due to the direct surface elevation signature of internal tides, which is of small amplitude, but to a shift in the barotropic tide induced by baroclinicity. Elevations are also more accurate in the two-layer model when pelagic tide gauges are used as the benchmark, and when the T/P-constrained TPXO6.2 model is used as a benchmark in deep waters south of 66°S. For Antarctic diurnal tides, the improvement in forward model elevation accuracy with baroclinicity is substantial. The optimal multiplicative factor in the two-layer case is nearly the same as in the one-layer case, against initial expectations that the explicit resolution of

*Corresponding author.
E-mail address: arbic@splash.princeton.edu (B.K. Arbic).

0967-0645/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dsr2.2004.09.014

low-mode conversion would allow less parameterized drag. In the optimally tuned two-layer M_2 solution, local values of the ratio of temporally averaged squared upper layer speed to squared lower layer speed often exceed 10.
© 2004 Elsevier Ltd. All rights reserved.

1. Introduction

Tides are the simplest large-scale motions in the open ocean. The barotropic tidal response in the mid-ocean is linear. Compared to the oceanic general circulation, tides are controlled by relatively few physical processes. In contrast to the wind- and buoyancy-forcing fields, tidal forcing has a spatial and temporal structure that is known very accurately. Yet it is only in the last few decades that open ocean tidal elevations have been mapped to high accuracy. Accurate charts have been constructed with models that are constrained by either tide-gauge or satellite altimetry data. The relative simplicity of tides makes the challenge of mapping them globally in forward models, that is, models that are unconstrained by data, an appealing one. Forward global tide models are an ideal testing ground for the hydrodynamical cores of numerical ocean general circulation models, and for ideas about drag and dissipation. In contrast to data-constrained models, forward models cannot achieve accurate tidal evelations unless substantial parameterized drag is included in the abyss. Forward models thus point clearly to drag in the open ocean as a central control on tidal flow. In agreement with other recent studies of global forward tide models (Jayne and St. Laurent, 2001—hereafter JS; Carrere and Lyard, 2003—hereafter CL; Egbert et al., 2004—hereafter ERB), we argue that topographic drag is a plausible mechanism for mid-ocean tidal dissipation. (Note that CL focused mostly on wind-driven motions and discussed tides only briefly.) Topographic drag probably affects all large-scale oceanic motions, and the understanding gained through use of forward tide models should be helpful in planning numerical simulations that simultaneously contain tides and the general circulation. This paper focuses on three factors—self-attraction and loading, topographic drag, and baroclinicity—that affect the accuracy of global forward modeled tidal elevations. To the best of our knowledge, the two-layer simulations presented in the current paper and the simulations of Simmons, Hallberg, and Arbic (this issue; hereafter SHA) are the first published global simulations of baroclinic tides. There are several differences between SHA and the present study. In place of the full spherical harmonic computation of the self-attraction and loading term, SHA used the scalar approximation, which facilitated the performance of higher resolution simulations. SHA focused on the rate of energy conversion from barotropic to resolved low-mode baroclinic motions, rather than the accuracy of surface elevations, and SHA did not utilize a parameterization of topographic drag.

Both observational and theoretical work on tides have a long history (Cartwright, 1999). Much of the theoretical work has been done in idealized geometries. The development of modern computers led to numerical tidal models with realistic geometry and bathymetry. Hendershott (1977, 1981) and Schwiderski (1980) review the pioneering work. Early successes at constraining hydrodynamical tide models by tide-gauge data were achieved by Parke and Hendershott (1980) and Schwiderski (1980). A tide model was derived solely from Geosat altimetric data by Cartwright and Ray (1990). Both the Schwiderski and Cartwright and Ray models have a combined elevation accuracy over the eight largest tidal constituents of about 4.5 cm, relative to deep-ocean tide gauges (Shum et al., 1997). With the launch of the TOPEX/POSEIDON (T/P) altimeter in 1992, the need to subtract out tidal elevations, so that other oceanic motions could be studied, stimulated the development of over 20 data-constrained models. All agreed with deep-ocean tide gauges to order 2.5–3 cm (Andersen et al., 1995; Le Provost et al., 1995; Shum et al., 1997). Le Provost et al. (1994) ran a hydrodynamical model which utilized tide-gauge data, albeit in a somewhat subtle manner—

see ERB for a discussion. Since the Le Provost et al. model was independent of altimetric data, it served as an important check on other data-constrained models, which were either entirely derived from T/P data (e.g. Desai and Wahr, 1995) or assimilated T/P data into hydrodynamical cores (e.g. Egbert et al., 1994).

The total dissipation of tidal energy can be inferred accurately from various astronomical methods (e.g., Munk, 1997 and references therein; Cartwright, 1993; Ray, 1994; Kagan and Sundermann, 1996; Munk and Wunsch, 1998). In the last 10 years or so, these methods have converged on 3.7 TW for the total tidal dissipation in the Earth–Moon–Sun system. An estimated 0.2 TW of dissipation occurs in the solid-earth tides, leaving 3.5 TW to be dissipated in the oceans, with M_2 accounting for 2.4 TW. The geographical distribution of tidal dissipation has been a matter of longstanding interest. Taylor (1919) estimated 0.04 TW of dissipation in the Irish Sea from an area integral of $c_d\rho_0\langle|\vec{u}|^3\rangle$, where brackets denote time-averaging, the bulk drag coefficient $c_d = 0.0025$, $\rho_0 = 1035\,\mathrm{kg\,m^{-3}}$ is the average density of seawater, and $\vec{u}$ is the horizontal velocity vector. Jeffreys (1920) applied the method globally and estimated 2 TW of dissipation in all of the world's shallow seas. Miller (1966) arrived at a similar total dissipation but a rather different partition amongst the various shallow seas. Shallow seas cover a more limited area than the deep ocean but have much larger tidal velocities. Abyssal dissipations based on the nominal c_d value of 0.0025 would be two orders of magnitude too small to have a significant impact on the global dissipation budget (Munk, 1997).

The tidal power input P_{in} can be shown to equal the global integral

$$P_{\mathrm{in}} = \rho_0 g \int\int \left\langle \eta_{\mathrm{EQ}} \frac{\partial \eta}{\partial t} \right\rangle \mathrm{d}A, \tag{1}$$

(e.g., Egbert and Ray, 2001 and references therein), where t is time, the gravitational acceleration $g = 9.80\,\mathrm{m\,s^{-2}}$, η_{EQ} is the equilibrium tide, to be discussed later, η is the surface tidal elevation with respect to mean resting water depths, and $\mathrm{d}A$ is an element of area. Since η_{EQ} is known to high accuracy and since data-constrained models have accurate elevations, at least in the open ocean, several recent data-constrained models (Le Provost and Lyard, 1997; Kantha, 1998; Tierney et al., 2000) have globally integrated power inputs (and, therefore, dissipation rates, since there must be a balance in the time-average) that are in good agreement with the astronomical constraints. The spatial distribution of the dissipation, however, is not constrained by the global integral. Bulk-drag friction in the models cited above puts virtually all of the modeled dissipation into shallow seas. On the other hand, in-situ observations at specific locations (e.g. Armi, 1978; Polzin et al., 1997; Lueck and Mudge, 1997; Kunze and Toole, 1997; Ledwell et al., 2000) suggest that substantial energy dissipation occurs over mid-ocean rough topography. Ledwell et al. (2000) showed that the magnitude of column-integrated energy dissipation in Brazil Basin microstructure measurements modulates with the spring-neap cycle, suggesting that much of the dissipation is driven specifically by tides. Egbert and Ray (2000, 2001, 2003a) infer from T/P-constrained models that about 0.8 TW of M_2 dissipation, and 1 TW of dissipation across all constituents, occurs in the deep oceans, primarily over regions of rough topography.

Tidal flow over rough topography generates high-mode internal waves which break and dissipate, and low-mode internal tides (Wunsch, 1975; Hendry, 1977; Hendershott, 1981) which radiate. Relevant numerical and theoretical studies of internal wave and internal tide generation over rough topography include Cox and Sandstrom (1962), Baines (1973), Bell (1975a), Bell (1975b), Baines (1982), Hibiya (1986), Sjoberg and Stigebrandt (1992), Morozof (1995), Holloway (1996), Cummins and Oey (1997), Kantha and Tierney (1997), Holloway and Merrifield (1999), Kang et al. (2000), Merrifield et al. (2001), Niwa and Hibiya (2001), Balmforth et al. (2002), St. Laurent and Garrett (2002), Llewellyn Smith and Young (2002, 2003), Khatiwala (2003), St. Laurent et al. (2003), Legg (2004), and Polzin (2004). Some of the above studies are done in idealized conditions (e.g., Khatiwala, 2003). Some combine theories with velocity outputs from one-layer data-constrained models to estimate global barotropic to baroclinic conversion rates (e.g., Nycander, submitted 2004).

Some are realistic regional simulations (e.g., Holloway, 1996; Cummins and Oey, 1997; Kang et al., 2000; Merrifield et al., 2001). Niwa and Hibiya (2001) modeled the internal tides of the entire North Pacific. Coherent radiating low-mode internal tides have been found in acoustic tomography data (Dushaw et al., 1995) and in T/P data (Ray and Mitchum, 1996, 1997; Cummins et al., 2001), and a comprehensive observational and modeling study of the internal tides at Hawaii is described in Rudnick et al. (2003).

Consistent with JS, CL, and ERB, we show that the accuracy of forward-modeled tidal elevations depends to first order on the amount of parameterized abyssal dissipation. We show that even physically implausible drag schemes, such as the introduction of inordinately large bulk quadratic drag or horizontal viscosity coefficients, can produce reasonably accurate tides. However, following JS, CL, and ERB, we argue that topographic drag schemes are more physically plausible. Our topographic drag scheme is based on an exact linear solution for arbitrary terrain, augmented by a dimensional analysis to account for nonlinear drag due to bottom turbulence as well as linear drag due to radiating waves. In contrast to those in JS, but similar to those in ERB, our solutions are convergent with respect to iterations of the self-attraction and loading (SAL) term, and therefore represent self-consistent forward solutions for the global tide. As in JS, we perform multi-constituent forward model runs, whereas ERB focused on M_2-only simulations. The regional and basin-wide studies of baroclinic tides previously noted have been done at high horizontal resolution (for instance, 1/10° in Kang et al., 2000; 4 km in Merrifield et al., 2001, and 1/16° in Niwa and Hibiya, 2001), and often include many vertical layers. We show here that even 1/2° horizontal resolution and two layers in the vertical are enough to provide a useful result—namely, an improvement in the accuracy of surface tidal elevations over those in our one-layer model. In agreement with Ray and Mitchum (1996, 1997), we find that the direct signature of baroclinic tides on the surface elevation is small. The improvement in surface elevation accuracy in the two-layer model is due to a shift in the barotropic solution.

Our main focus is on tides in the open ocean. Coastal tides have smaller spatial scales than open ocean tides, are more nonlinear, and exhibit aperiodic behaviors (Maas and Doelman, 2002). Obtaining accurate coastal tides in a global model is a formidable challenge which even data-constrained models are only beginning to meet (Ray, 1999; Tierney et al., 2000). We follow JS in referring to regions having resting water column depths less than 1000 m as "shallow" and regions having depths greater than 1000 m as "deep", "abyssal", "open-ocean", or "mid-ocean". The T/P-constrained model output we compare our forward model to was provided courtesy of R.D. Ray. We draw the four largest semidiurnal constituents (N_2, M_2, S_2, and K_2) and four largest diurnal constituents (Q_1, O_1, P_1, and K_1) from the GOT99.2 model (Ray, 1999), which is an empirical mapping guided by the Le Provost et al. (1994) results. Purely empirical derivations from T/P, documented in Egbert and Ray (2003b, see their Fig. 1a,b) provide us with the two largest long-period constituents (M_m and M_f). As with other recent forward tide models, our forward model contains one free parameter, associated with the topographic drag scheme. We tune this parameter to minimize the surface elevation discrepancy in M_2-only runs with respect to GOT99.2, in deep waters over the latitudes covered by the T/P altimeter (66°S to 66°N). The optimal value of the free parameter so chosen is then used in multi-constituent runs. We compare the multi-constituent forward model elevations to those in the Ray results and to those in the set of 102 pelagic tide gauges used in Shum et al. (1997). Another test of our forward model can be made in latitudes poleward of 66°. Padman et al. (2002, 2003) and references therein present Antarctic tide models constrained by tide-gauge data, and Padman and Erofeeva (2004) and references therein present tide-gauge constrained Arctic tide models. Elevations in the global TPXO6.2 solution (Egbert and Erofeeva, 2002) compare well with Antarctic tide-gauge data (Laurie Padman, Gary Egbert, and Lana Erofeeva, personal communication 2004). We will therefore use TPXO6.2 as a benchmark in latitudes south of 66°S. The comparisons to pelagic tide gauges and to TPXO6.2 do not include

M_f and M_m. Against all three benchmarks, elevations in our forward two-layer model are more accurate than those in our one-layer model. Parameterized drag is needed to obtain accurate tides in both one- and two-layer simulations, and we provide a simple explanation for the sensitivity of elevation accuracy to drag. Although the main focus of this paper is on surface elevations, we also devote a brief section to the impact of baroclinicity on tidal velocities.

2. The model

2.1. Governing equations

We have adapted the Hallberg Isopycnal Model (HIM; Hallberg and Rhines, 1996) to be a forward model of tides. Our simulations are performed on a latitude–longitude grid running from 86°S to 82°N, which allows the possibility of accurate tidal modeling in the Antarctic, but not in the high Arctic. (Note that SHA utilizes a tripolar grid to perform truly global simulations.) Except where noted, the runs discussed in this paper are at 1/2° resolution. A few runs are at 1/4°. The one-layer mass conservation equation is

$$\frac{\partial \eta}{\partial t} + \nabla \cdot [(H+\eta)\vec{u}] = 0, \tag{2}$$

where H is the resting layer thickness (spatially varying, according to the topography). The one-layer momentum equation is

$$\begin{aligned}\frac{\partial \vec{u}}{\partial t} &+ (f+\zeta)\hat{k}\times\vec{u} \\ &= -g\nabla(\eta-\eta_{EQ}-\eta_{SAL}-\eta_{MEM}) - \nabla\left(\frac{1}{2}\vec{u}\cdot\vec{u}\right) \\ &+ \frac{\nabla\cdot[K_H(H+\eta)\nabla\vec{u}]}{H+\eta} - \frac{c_d|\vec{u}|\vec{u}}{H+\eta} + \frac{\overline{T}\vec{u}}{\rho_0(H+\eta)},\end{aligned} \tag{3}$$

where f is the Coriolis parameter, $\hat{k}$ is a unit vector in the vertical direction, $\zeta = \hat{k}\cdot(\nabla\times\vec{u})$ is the vorticity, η_{EQ} is the equilibrium tidal forcing, η_{SAL} is the self-attraction and loading term, η_{MEM} is a "memory" term we have inserted which we discuss later, c_d is set to the canonical value of 0.0025 unless otherwise noted, K_H is a horizontal viscosity set to $10^3\,m^2\,s^{-1}$ unless otherwise noted, and $\overline{T}$ is our topographic drag tensor, to be discussed later. The horizontal viscosity prevents sub-grid scale noise brought about by the non-linear advection terms in the governing equations (Egbert and Ray, 2003b), and prevents slow drifts in model energy. The chosen value of horizontal viscosity affects modeled energy dissipations at the level of a few percent or less.

The two-layer mass conservation equations are

$$\frac{\partial(\eta_1-\eta_2)}{\partial t} + \nabla\cdot[(H_1+\eta_1-\eta_2)\vec{u}_1] = 0, \tag{4}$$

$$\frac{\partial \eta_2}{\partial t} + \nabla\cdot[(H_2+\eta_2)\vec{u}_2] = 0, \tag{5}$$

and the momentum equations are

$$\begin{aligned}\frac{\partial \vec{u}_1}{\partial t} &+ (f+\zeta_1)\hat{k}\times\vec{u}_1 \\ &= -g\nabla(\eta_1-\eta_{EQ}-\eta_{SAL}-\eta_{MEM}) \\ &- \nabla\left(\frac{1}{2}\vec{u}_1\cdot\vec{u}_1\right) + \frac{\nabla\cdot[K_H(H_1+\eta_1-\eta_2)\nabla\vec{u}_1]}{H_1+\eta_1-\eta_2},\end{aligned} \tag{6}$$

$$\begin{aligned}\frac{\partial \vec{u}_2}{\partial t} &+ (f+\zeta_2)\hat{k}\times\vec{u}_2 \\ &= -g\nabla(\eta_1-\eta_{EQ}-\eta_{SAL}-\eta_{MEM}) \\ &- g'\nabla(\eta_2-\eta_{EQ}-\eta_{SAL}-\eta_{MEM}) \\ &- \nabla\left(\frac{1}{2}\vec{u}_2\cdot\vec{u}_2\right) + \frac{\nabla\cdot[K_H(H_2+\eta_2)\nabla\vec{u}_2]}{H_2+\eta_2} \\ &- \frac{c_d|\vec{u}_2|\vec{u}_2}{H_2+\eta_2} + \frac{\overline{T}\vec{u}_2}{\rho_0(H_2+\eta_2)},\end{aligned} \tag{7}$$

where subscripts 1 and 2 denote upper and lower layers, respectively, H_1 and H_2 are the resting layer depths, η_1 and η_2 are the surface and interfacial height perturbations, and the reduced gravity $g' = (\rho_2-\rho_1)/\rho_0$, where ρ_1 and ρ_2 are the layer densities. Note that we have chosen to have the drag tensor operating on the lower-layer flow.

Many references exist on the equilibrium tide η_{EQ} (e.g., Cartwright, 1977; Hendershott, 1981; Marchuk and Kagan, 1984; Pugh, 1987), which is astronomically forced but modified by a factor of $1+k_2-h_2$, where the Love numbers h_2 and k_2

Table 1
Constituent-dependent frequencies ω, astronomical forcing amplitudes A, and Love number factors $1+k_2-h_2$ used to compute equilibrium tide η_{EQ}. The periods $2\pi/\omega$ are also given

Constituent	ω (10^{-4} s^{-1})	A (cm)	$1+k_2-h_2$	Period (solar days)
M_m	0.026392	2.2191	0.693	27.5546
M_f	0.053234	4.2041	0.693	13.6608
Q_1	0.6495854	1.9273	0.695	1.1195
O_1	0.6759774	10.0661	0.695	1.0758
P_1	0.7252295	4.6848	0.706	1.0027
K_1	0.7292117	14.1565	0.736	0.9973
N_2	1.378797	4.6397	0.693	0.5274
M_2	1.405189	24.2334	0.693	0.5175
S_2	1.454441	11.2743	0.693	0.5000
K_2	1.458423	3.0684	0.693	0.4986

respectively account for the solid-earth body-tide deformation and the perturbation gravitational potential resulting from this deformation (Hendershott, 1972; Ray, 1998). Table 1, taken from documentation for the TPXO6.2 model (Egbert and Erofeeva, 2002), lists the frequencies, amplitudes, and Love numbers for the 10 constituents simulated in this paper. The frequencies and amplitudes are based on work by Doodson (1921), Cartwright and Tayler (1971), and Cartwright and Edden (1973)—a more recent computation was done by Roosbeek (1996). The Love numbers for the diurnal tides differ from those for the semi-diurnal and long-period tides because of the free-core nutation resonance (Wahr, 1981; Wahr and Sasao, 1981). When all 10 constituents are present, the equilibrium tide is

$$\begin{aligned}\eta_{EQ}(\phi,\lambda,t) &= \sum_l A(l)[1+k_2(l)-h_2(l)] \\ &\times\left(\frac{1}{2}-\frac{3}{2}\sin^2\phi\right)\cos[\omega(l)t] \\ &+\sum_d A(d)[1+k_2(d)-h_2(d)] \\ &\times\sin(2\phi)\cos[\omega(d)t+\lambda] \\ &+\sum_s A(s)[1+k_2(s)-h_2(s)]\cos^2\phi \\ &\times\cos[\omega(s)t+2\lambda], \end{aligned} \tag{8}$$

where l is summed over M_m and M_f, d is summed over the four diurnal tides, s is summed over the four semi-diurnal tides, λ is longitude with respect to the Greenwich meridian, and ϕ is latitude. We set the time t to zero at the start of each numerical simulation. We ignore absolute time and slow modulations of phase or amplitude, all of which must be considered in more precise tidal analysis and prediction (e.g. Schwiderski, 1980; Pugh, 1987).

Most of our runs are M_2-only or K_1-only. In these cases we run the model for 20 days (equilibration typically takes about 12 days). All quantities discussed in this paper are drawn from the model after equilibration has been reached. Time averaged quantities are computed over a period of duration one forcing cycle, sampled every 12 min. Simulations of long-period tides require longer integrations. Frequencies are separated from each other in multi-constituent runs with the harmonic analysis package of Pawlowicz et al. (2002), which is based on that of Foreman (1977). Satisfaction of the Rayleigh criterion for separation of the K_2/S_2 and K_1/P_1 pairs requires an analysis period at least 182 days long. We run our multi-constituent simulations for 240 days, discard the first 40, and perform the harmonic analysis on the last 200. Because multi-constituent runs are time-consuming and computationally expensive, we perform only a few of them in this paper, after initial explorations have been done in single-constituent runs.

The perturbation potential due to the gravitational self-attraction of the ocean surface is

$$\sum_n \frac{3g\rho_0}{\rho_{\text{earth}}} \frac{\eta_n}{2n+1}, \tag{9}$$

where $\rho_{\text{earth}} = 5518\,\text{kg m}^{-3}$ is the average density of the solid earth and the η_n's are order n spherical harmonics of the tidal elevation η (Hendershott, 1972; Ray, 1998). The solid earth yields to the load of the water column, and the gravitational potential is thus altered by the resulting redistribution of mass within the solid earth. The synthesis of these arguments is

$$\eta_{\text{SAL}} = \sum_n \frac{3\rho_0}{\rho_{\text{earth}}(2n+1)}(1 + k'_n - h'_n)\eta_n \tag{10}$$

(Hendershott, 1972; Ray, 1998). The load numbers h'_n and k'_n, introduced in Munk and MacDonald (1960), respectively account for solid-earth yielding and the resulting perturbation potential. We take the load numbers from Table A2 in Farrell (1972), who calculated them from a spherically symmetric solid-earth model. For our 1/2° model, n goes from 1 to 360. We use a spline fit to fill in the gaps in Farrell's table. Usage of load numbers from more recent, but still spherically symmetric, solid-earth models (Jerry Mitrovica, 2003, personal communication) makes little difference to our tide model results.

Calculation of spherical harmonics as the numerical tide model is running is computationally infeasible. The simplest alternative is to use the "scalar approximation"

$$\eta_{\text{SAL}} \approx \beta\eta, \tag{11}$$

where β is a constant, often taken to be 0.085 (Accad and Pekeris, 1978). We estimate β by computing a least-squares fit in depths greater than 1000 m to η and η_{SAL} values in the GOT99 model. A representative value for the semi-diurnal tides is $\beta = 0.094$, and we use that as our nominal value. In some runs (i.e. those in Section 4.2) we use the scalar approximation and proceed no further. In some we begin with the scalar approximation and then employ the full formula (10) in an iterative procedure (Francis and Mazzega, 1990). In Section 4.1 we describe this procedure and the term η_{MEM}, which is introduced into the momentum equation to force convergence of η_{SAL} (we set η_{MEM} to zero in all runs which employ the scalar approximation).

2.2. Diagnostics

The main diagnostic we use to evaluate model accuracy is the sea-surface height discrepancy D against benchmark elevations $\eta_{\text{BENCHMARK}}$, where the latter are drawn variously from GOT99.2, TPXO6.2, and the set of 102 pelagic tide gauges used in Shum et al. (1997):

$$D = \sqrt{\frac{\langle \iint (\eta - \eta_{\text{BENCHMARK}})^2 \, dA \rangle}{\iint dA}}. \tag{12}$$

Throughout the main body of this paper, brackets refer to time-averaging. When η and $\eta_{\text{BENCHMARK}}$ refer to the combined elevations of several constituents, we refer to D as the RSS (root-sum-square) discrepancy, as is customary in the literature. By computing the signal

$$S = \sqrt{\frac{\langle \iint (\eta_{\text{BENCHMARK}})^2 \, dA \rangle}{\iint dA}}, \tag{13}$$

we can compute relative errors, for instance, the percentage of sea surface height variance captured, given by $100 * [1 - (D/S)^2]$. As with D, S can be calculated either for individual constituents or for their (RSS) combination. Our calculations of D and S which utilize GOT99.2 or TPXO6.2 as benchmarks are restricted to deep waters (resting depths exceeding 1000 m). We also restrict most calculations of D and S to latitudes equatorward of 66°; exceptions will be noted. In the pelagic tide gauge comparisons, as is customary, we compute averages that are not weighted by area or any other considerations.

The available potential energy (APE) and kinetic energy (KE) are given by

$$KE = \frac{1}{2}\rho_0 \iint (H + \eta)\vec{u} \cdot \vec{u} \, dA,$$

$$APE = \frac{1}{2}\rho_0 g \iint \eta^2 \, dA, \tag{14}$$

in the one-layer case, and by

$$KE = \frac{1}{2}\int\int [\rho_1(H_1+\eta_1-\eta_2)\vec{u}_1 \cdot \vec{u}_1 + \rho_2(H_2+\eta_2)\vec{u}_2 \cdot \vec{u}_2]\,dA,$$
$$APE = \frac{1}{2}\rho_0 \int\int (g\eta_1^2 + g'\eta_2^2)\,dA, \quad (15)$$

in the two-layer case. The kinetic energy can also be written as the sum of barotropic (BT) and baroclinic (BC) components $KE = KE_{BT} + KE_{BC}$, where

$$KE_{BT} = \frac{1}{2}\rho_0 \int\int (H_1+H_2+\eta_1)\vec{u}_{BT} \cdot \vec{u}_{BT}\,dA,$$

$$KE_{BC} = \frac{1}{2}\rho_0 \int\int (H_1+H_2+\eta_1)\vec{u}_{BC} \cdot \vec{u}_{BC}\,dA,$$

$$\vec{u}_{BT} = [(H_1+\eta_1-\eta_2)\vec{u}_1 + (H_2+\eta_2)\vec{u}_2]/(H_1+H_2+\eta_1),$$

$$\vec{u}_{BC} = \sqrt{(H_1+\eta_1-\eta_2)(H_2+\eta_2)} \times(\vec{u}_1-\vec{u}_2)/(H_1+H_2+\eta_1). \quad (16)$$

In the latter case the second layer will not exist in waters shallower than $H_{\text{initupper}} - \eta_2$, where $H_{\text{initupper}}$ is the initial value of the upper layer thickness.

2.3. Input datasets

Our topographic drag scheme requires the Brunt–Vaisala frequency N at the seafloor, which we compute from the Levitus et al. (1998) climatology. Our bathymetry comes from a 2 min dataset assembled by Andrew Coward. Between 72°S and 72°N, Coward (2004, personal communication) mapped version 6.2 of the Smith and Sandwell (1997) dataset from the original Mercator projection onto a latitude–longitude grid by choosing median values over each model grid cell, without performing any smoothing. North of 72°N, an early version of the IBCAO (Jakobssen et al., 2000) dataset was used, without any smoothing at the joins, while ETOPO5 was joined without smoothing to the dataset south of 72°S. In waters that are south of 60°S and shallower than 1200 m, we alter Coward's dataset with data obtained from Laurie Padman (2002, persona communication). The appendix in Padman et al (2002) describes the sources of Antarctic bathymetry data. Under the Antarctic ice shelves, water depths are measured from the seafloor to the bottom of the ice.

The 1/30° topography $h_{\text{fine}}(x', y')$ is filtered onto the 1/2° and 1/4° grid values $H(x, y)$ via

$$H(x,y) = \frac{\int\int h_{\text{fine}}(x',y')F(x,y,x',y')\,dA}{\int\int F(x,y,x',y')\,dA}. \quad (17)$$

We use a radial Blackman filter

$$F(x,y,x',y') = 0.42 + 0.5\cos(\pi r_{\text{field}}/r_{\text{filter}}) + 0.08\cos(2\pi r_{\text{field}}/r_{\text{filter}}), \quad (18)$$

where $r_{\text{field}} = \sqrt{(x-x')^2+(y-y')^2}$ is the distance from the coarse grid model gridpoint (x, y) to the field point (x', y'), and r_{filter} is the filter radius. For points lying outside the filter radius, $F = 0$. After some initial explorations, we chose a filter radius of 55 km (1/2° in latitude) to prepare both the 1/2° and 1/4° model grids used in this paper. Thus both grids have the same amount of roughness. The chosen 55 km filter radius does not satisfy the Nyquist criteria for the 1/2° simulations, which therefore contain noticeably more gridscale noise than the 1/4° simulations, in both the surface and interfacial elevation fields. However, in both the one- and two-layer runs, elevation accuracy is nearly equal in 1/2° as 1/4° simulations. The model elevation accuracy is sensitive to the type (e.g., Gaussian versus Blackman) and radius of the filter used to map the topography onto our model grid. The sensitivity to the preparation of the topographic grid, and the related sensitivity to resolution, will be further explored in a future paper.

3. Summary of topographic drag scheme

The classic topographic drag problem has a steady flow $\vec{u}$ over small-amplitude topography $h\sin(kx)$, where h and k are constants, the fluid has constant Brunt–Vaisala frequency N, and the internal wave drag is $\rho_0 N k h^2 \vec{u}$. Actual ocean topography contains a spectrum of heights and wavenumbers. For their forward tidal model, JS

first calculated the water depths H at every gridpoint in their 1/2° domain by fitting a polynomial through all the 1/30° Smith and Sandwell (1997) data contained by the coarse model gridpoint. The rms of the residuals of the nearby 1/30° heights with respect to the local H is taken to be the roughness h. JS added a factor of $\frac{1}{2}$ to the formula above, and then tuned k to minimize the misfit between their forward modeled elevations and those of data-constrained models.

In this section we summarize our topographic drag scheme (Garner, 2003; also, submitted 2004), which was originally developed for quasi-steady background flows in the atmosphere, and then extended to the oscillatory case relevant for tides. Details of the scheme are provided in the Appendix. Instead of using statistical measures of the various vertical and horizontal scales in the topography, the scheme is based on an exact analytical solution for drag in the hydrostatic limit over arbitrary small-amplitude terrain $h(\vec{x})$. The drag in the oscillatory case depends on tidal frequency. Since M_2 is the dominant tidal constituent, we tune our drag scheme in M_2-only runs and use the resulting optimal drag parameter in the drag formula for our multi-constituent runs. Our application of the drag scheme is therefore not strictly correct for the diurnal and long-period tides in such runs. However, the accuracy of diurnal tidal elevations is not as sensitive to drag as is the accuracy of semi-diurnal elevations, and the accuracy of long-period elevations exhibits very little sensitivity.

Inspired by the meteorological literature on topographic drag and drag parameterizations (e.g., Pierrehumbert, 1987; Miranda and James, 1992; Lott and Miller, 1997), we employ dimensional reasoning in our scheme to accommodate the possibility of drag due to low-level wave breaking and turbulence as well as that due to linear radiating waves. Waves radiating away from small mountains or the tips of tall mountains cause drag proportional to $|\vec{u}|$, while low-level breaking in flow deflected around tall mountains causes drag proportional to $|\vec{u}|^2$. Flow splitting occurs when a nondimensional parameter exceeds an order one critical value. The measure of nonlinearity depends on whether or not the topographic length scales considered are small enough that oscillatory flows appear to be "steady". The nonlinearity measure for large scales is small, indicating that flow splitting and nonlinear drag probably occur mostly at small scales. The nonlinearity parameter for small scales is the Froude number $Nh/|\vec{u}|$. Thurnherr and Richards (2001), Thurnherr et al. (2002), and Thurnherr and Speer (2003) argue from observational data that turbulence associated with hydraulic jumps occurs in areas of rough topography when the Froude number exceeds an order one threshold. The end products of our dimensional reasoning are a "propagating" drag D_p due to radiating waves (which presumably break above the topography) and a "non-propagating" drag D_{np} due to low-level turbulence. Our scheme takes the drag direction from the linear analysis, and matches the small amplitude limit of D_p to the linear analytical result. The final result for our drag is given in the Appendix.

In all of the runs presented in this paper, our drag scheme operates in waters 1000 m and deeper. JS used their scheme in waters 100 m and deeper. Preliminary investigations indicate that the accuracy of our forward-modeled tidal elevations improves slightly if we use our drag scheme in waters 100 m and deeper, and our future work may implement this. The horizontal length scale over which to compute the quantity χ defined in the Appendix (Eq. (31)) is another ambiguous parameter. Larger integration scales lead to larger values of χ and thus drag. We choose to integrate (31) over 40 km, which is an order-one fraction of the order 100 km topographic length scales involved in generating the first-mode baroclinic tide (see Fig. 11 in SHA). Topographic scales much larger than this do not generate internal waves and should not be included.

Much of the turbulent kinetic energy dissipation seen in microstructure data (Lou St. Laurent, 2003, personal communication) occurs within 500 m of the bottom. Internal waves with vertical scales much smaller than this are generated by horizontal topographic scales of order hundreds of meters to a few kilometers. Although the Smith and Sandwell (1997) topographic dataset is mapped onto a 1/30° grid, its true horizontal

resolution is of order π times the ocean depth, i.e. about 12 km. Thus the Smith and Sandwell dataset is missing the horizontal scales that generate breaking high-mode internal waves (and low-level turbulence). We have found that our estimated topographic drag must be augmented by a multiplicative factor to obtain accurate tidal elevations. We tune the multiplicative factor to minimize D. Our present interpretation is that the multiplicative factor makes up for the missing scales in the topographic dataset, but we will examine this assumption more closely in future work. We stress again that JS also had a tunable parameter (k). We show that both schemes, when optimally tuned, produce nearly identical overall levels of drag, as measured by the spatially-averaged decay rate. In our one-layer runs, the spatially averaged decay rate is

$$r_{\text{average}} = \frac{\iint \left(\left| \frac{\overline{\tau}\vec{u}}{\rho_0(H+\eta)} \right| / |\vec{u}| \right) \mathrm{d}A}{\iint \mathrm{d}A}. \tag{19}$$

For the JS scheme the spatially averaged decay rate is

$$r_{\text{average}} = \frac{\iint (kh^2 N/2H)\, \mathrm{d}A}{\iint \mathrm{d}A}. \tag{20}$$

4. Surface elevation accuracy of M_2-only simulations

This section explores the surface elevation accuracy in M_2-only simulations. We examine the effects of self-attraction and loading in one-layer runs, and the effects of parameterized drag in both one- and two-layer runs. The results of the exploration are utilized in the multi-constituent runs of the next section.

4.1. Effects of self-attraction and loading

Although the amplitude of η_{SAL} is only about one-tenth that of η, $\nabla\eta_{\text{SAL}}$ is an order-one perturbation in the momentum equation because η_{SAL} contains much smaller scales than does η_{EQ} (Hendershott, 1972; Gordeev et al., 1977). To confirm the critical importance of this term, we performed 1/2° M_2-only simulations which completely omit η_{SAL}. Our topographic drag scheme was used, augmented by a multiplicative factor tuned as described in the next subsection. The minimum D value was 19.32 cm, much larger than that obtained when the scalar approximation (11) or the full expression (10) is used for η_{SAL}. In K_1-only simulations which omit the η_{SAL} term, the elevation errors against GOT99.2 were also substantially larger than in simulations which employ the scalar approximation, although the difference between the two cases is not as dramatic as it is with M_2.

We now describe our procedure for iterating η_{SAL} past the scalar approximation, in 1/2° runs which employ an optimally tuned topographic drag scheme. Once the first iteration, employing the scalar approximation, is completed, η_{SAL} is calculated offline from the model output via (10). Amplitude and phase maps $A_{\text{SAL}}(\phi,\lambda)$ and $p_{\text{SAL}}(\phi,\lambda)$ are created which satisfy

$$\eta_{\text{SAL}}(\phi,\lambda,t) = A_{\text{SAL}}(\phi,\lambda)\cos[\omega t - p_{\text{SAL}}(\phi,\lambda)]. \tag{21}$$

As is standard, the phase is referenced to the Greenwich meridian. The next iteration of the model computes η_{SAL} from the amplitude and phase maps made from the first iteration. When the new iteration is complete, new maps (21) are again created offline for use in the following iteration. Convergence can be obtained with this straightforward iteration technique in K_1-only simulations, but in M_2-only runs the value of D oscillates (solid curve in Fig. 1). Similar oscillations were seen by JS, who consequently averaged their model results over the iterations (Steve Jayne, 2003, personal communication). We have been able to suppress the oscillations by employing "memory" of the tidal elevations in previous iterations. Let us denote the tidal elevations of the iteration just before the current iteration by η_{PREVIOUS}. We construct amplitude maps $A_{\text{PREVIOUS}}(\phi,\lambda)$ and phase maps $p_{\text{PREVIOUS}}(\phi,\lambda)$ for this quantity and store them for use in the current iteration:

$$\eta_{\text{PREVIOUS}}(\phi,\lambda,t) = A_{\text{PREVIOUS}}(\phi,\lambda) \times \cos[\omega t - p_{\text{PREVIOUS}}(\phi,\lambda)]. \tag{22}$$

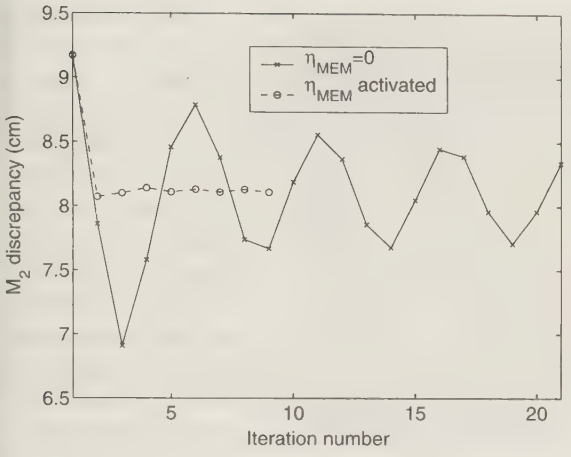


Fig. 1. Time- and area-averaged sea-level discrepancy D against GOT99.2 in iterations of 1/2° one-layer M_2-only runs, done with and without η_{MEM}. Discrepancies are calculated in waters deeper than 1000 m and equatorward of 66°. Iterations are with respect to η_{SAL}, the self-attraction and loading term.

For each iteration, we set η_{MEM} to be 0.094 (the same constant used in the scalar approximation) times the departure of the tide η in the current iteration from the elevation $\eta_{PREVIOUS}$ in the previous iteration:

$$\eta_{MEM} = 0.094(\eta - \eta_{PREVIOUS}). \quad (23)$$

With this technique, the sea-surface height discrepancy converges quickly (dashed curve in Fig. 1). Once the solution has converged, η_{MEM} is by definition a small term. Indeed, we found that when we continued the converged series plotted in Fig. 1, but with the η_{MEM} term removed, the discrepancies remain at the same level. Therefore it is not necessary to retain η_{MEM} past the first few iterations. ERB (see also references therein) also found that their iterations of η_{SAL} did not converge, unless an equivalent modified iteration procedure was employed.

The elevation error in the third iteration of our $\eta_{MEM} = 0$ set is substantially lower than that in the other iterations including all iterations in the convergent set. However, since the error does not remain at this low level, the third iteration in the $\eta_{MEM} = 0$ set is not a self-consistent solution. We believe therefore that the low error should be considered fortuitous. Using values of β other than 0.094 in (11) can also yield solutions with low elevation discrepancies. However, in these latter cases the discrepancy increases once η_{SAL} is iterated using the proper formula, and we believe the low discrepancies should again be considered fortuitous. In succeeding sections, whenever the self-attraction and loading term is iterated beyond the scalar approximation, η_{MEM} is employed, and the solutions are self-consistent.

4.2. *Effects of parameterized drag in one-layer runs*

This subsection explores the effects of parameterized drag on the accuracy of tidal elevations in one-layer 1/2° M_2-only experiments which, except where noted, employ the scalar approximation $\eta_{SAL} = 0.094\eta$ for simplicity. Even physically implausible drag can improve forward model accuracy. Fig. 2(A) plots the time-averaged global APE and KE in simulations which have no topographic drag, but in which c_d is treated as a tunable parameter. Both APE and KE decrease as c_d is increased. Fig. 2(B) shows the sea-surface height discrepancy D of the same experiments with respect to GOT99.2, computed in waters deeper than 1000 m and over latitudes equatorward of 66°. Since APE is the global integral of η^2, a quantity clearly related to D, we expect experiments with APE near that of observations to have relatively small elevation errors. Inspection of Fig. 2 indicates that this is the case. Thus drag affects surface elevation accuracy simply because it controls tidal amplitudes. The optimal experiment in Fig. 2 has a c_d value of 0.64, 256 times larger than the nominal value, which places significant dissipation (1.56 TW; see SHA for details of the model energy budget analysis) into the abyss and yields a total M_2 dissipation across deep and shallow waters of 3.09 TW, somewhat higher than the observed 2.44 TW value. Fig. 3 displays $\sqrt{\langle(\eta - \eta_{GOT99.2})^2\rangle}$, the spatial pattern of the temporally averaged M_2 elevation discrepancies, for both the nominal and optimal c_d experiments. Both experiments shown in this figure were iterated with respect to the self-attraction and loading term via the procedure described in Section 4.1. The amount of drag clearly affects

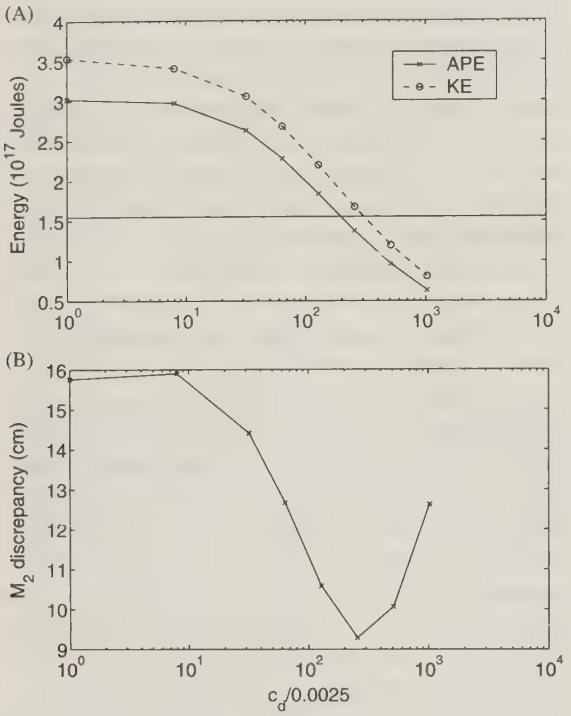


Fig. 2. (A) Global APE and KE in 1/2° one-layer M_2-only simulations with $\eta_{SAL} = 0.094\eta$ and variable c_d. The extra horizontal line represents the observed value of APE, taken from Tierney et al. (2000). (B) Time- and area-averaged sea-level discrepancy D against GOT99.2, in waters deeper than 1000 m and equatorward of 66°.

the spatial distribution as well as the magnitude of error. APE and D are similarly sensitive to the value of K_H in experiments (not shown) which treat it as a tunable parameter while omitting topographic drag. The optimal K_H in terms of producing accurate elevations is $4.5 \times 10^5\,m^2\,s^{-1}$. Again, this is at least two orders of magnitude larger than any physically plausible value.

The tunable c_d and tunable K_H simulations described above demonstrate that tidal amplitudes and thus accuracies are sensitive to drag, even if this drag is physically implausible. Fig. 4 shows that the same principles hold when we tune the multiplicative factor augmenting our topographic drag scheme, which is physically motivated, and concentrates parameterized dissipation into regions of rough topography (Fig. 5), in accordance with in situ (Polzin et al., 1997) and satellite (Egbert and Ray, 2000, 2001, 2003a) observations. Fig. 6 shows the spatial pattern of the elevation discrepancies against GOT99.2 in the optimally tuned one-layer M_2-only simulation. Again, as in Fig. 3, the simulations shown in Figs. 5 and 6 were iterated with respect to η_{SAL}. The optimal multiplicative factor for the iterated solutions is 7, lower than the optimal factor of 9 in runs employing the scalar approximation. Although the globally averaged elevation error is not much lower than that in the $c_d = 0.64$ run, the spatial distribution of error is different. The global tidal solution is affected by the spatially distributed nature of the topographic drag as well as its total magnitude.

4.3. *Effects of parameterized drag in two-layer runs*

To match the stratification of the ocean most closely in a two-layer simulation, one might choose a target density surface near the bottom of the thermocline and initialize its depths in the model according to a climatology. However, since such an interface would slope in the horizontal, it would undergo geostrophic adjustment in the absence of forcing which could maintain it. For simplicity, we instead choose to initialize our two-layer run with an interface placed uniformly at 700 m throughout the model domain. For similar reasons, we choose one representative value of g', $1.64 \times 10^{-2}\,m\,s^{-2}$, estimated for realistic subtropical gyre conditions from a formula in Flierl (1978) that depends on the chosen 700 m interfacial depth. We found that the elevation discrepancies D are not strongly sensitive to the exact value of g', which we varied from $1.2 \times 10^{-2}\,m\,s^{-2}$ to $1.8 \times 10^{-2}\,m\,s^{-2}$, or to the exact placement of the interfacial depth, which we varied from 600 to 900 m. Visual inspection of the interfacial waves indicates that they are marginally resolved in 1/2° runs, especially with smaller g' values, which render smaller horizontal scales. Two-layer 1/4° M_2-only experiments resolve the interfacial waves much better but yield nearly identical elevation errors.

It is an interesting question whether our drag parameterization should be thought of as representing conversion to low baroclinic modes or to high ones. This question takes on particular interest in our two-layer simulations, in which we

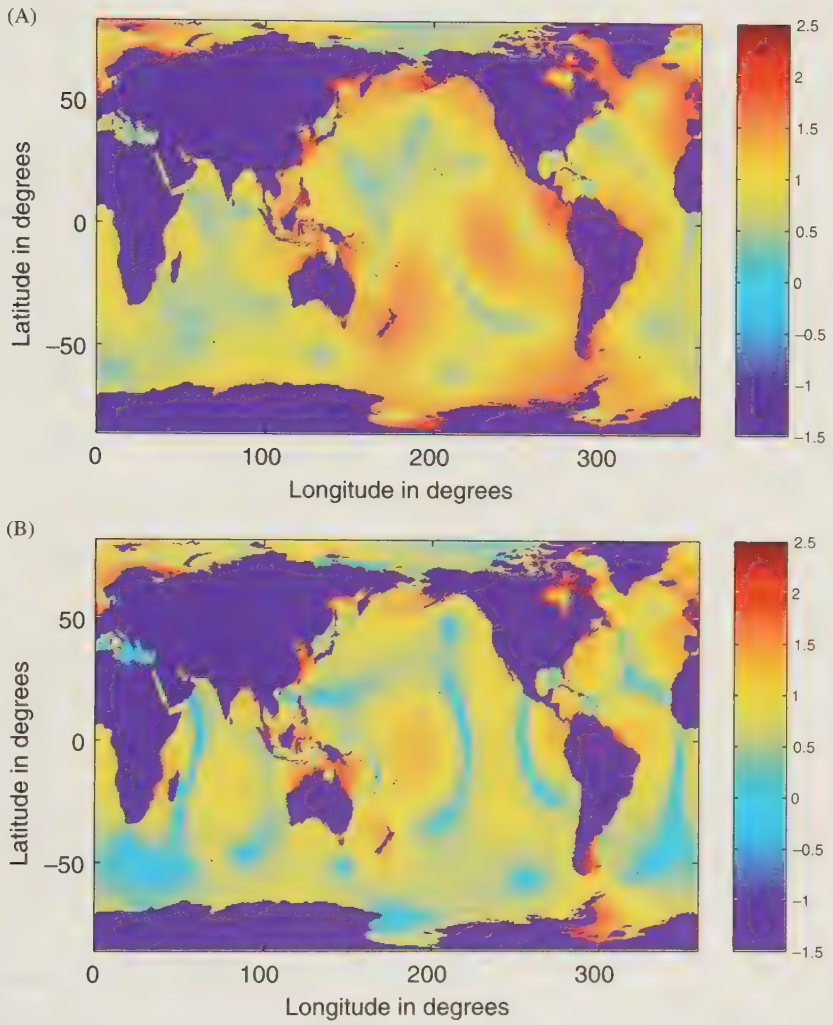


Fig. 3. Log_{10} of time-averaged sea level errors $\sqrt{\langle(\eta - \eta_{\mathrm{GOT99.2}})^2\rangle}$ (in cm) for 1/2° one-layer M_2-only: (A) $c_d = 0.0025$ experiment and (B) $c_d = 0.64$ experiment.

resolve the generation of the first baroclinic mode internal tide. Our initial expectation therefore was that less parameterized drag would be required to obtain accurate surface elevations in two-layer experiments. However, the sensitivity of two-layer M_2 simulations to the multiplicative factor is similar to the sensitivity in one-layer experiments. Two-layer M_2-only, 1/4° experiments having no parameterized topographic drag (i.e., only a c_d value of 0.0025 for dissipation) yield an rms elevation error of 12.8 cm (against GOT99.2, in waters deeper than 1000 m and equatorward of 66°) when the full iteration of self-attraction and loading is used. This represents an improvement over the 14.9 cm error we obtain in a nominal c_d-only one-layer run, but is much larger than the 7.4 cm error obtained with an optimally tuned multiplicative factor. The optimal multiplicative factor of 8 is not much different from the optimal factor of 7 found in one-layer simulations. Our preliminary interpretation of this fact is as follows. As was shown in the previous subsection, energy must be taken out of the system, via parameterized drag in the abyss, in order to bring about accurate

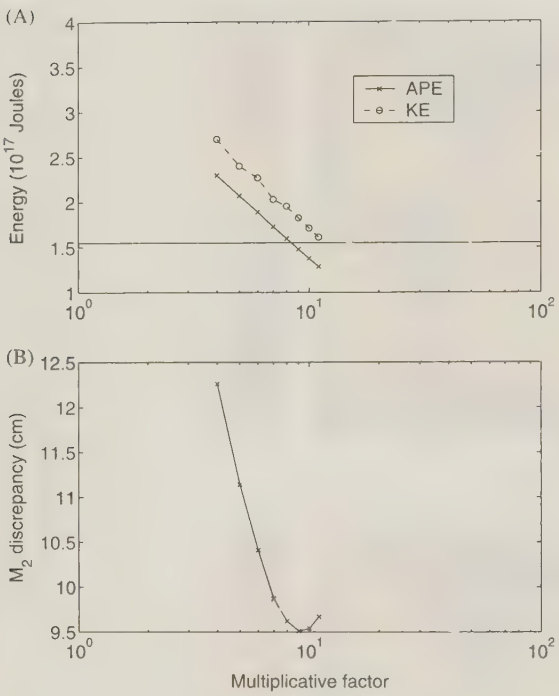


Fig. 4. (A) Global APE and KE in 1/2° one-layer M_2-only simulations with $\eta_{SAL} = 0.094\eta$, run with our topographic drag scheme augmented by a variable multiplicative factor. The extra horizontal line represents the observed value of APE, taken from Tierney et al. (2000). (B) Time- and area-averaged sea-level discrepancy D against GOT99.2, in waters deeper than 1000 m and equatorward of 66°.

tidal energies and elevations. In two-layer experiments, a substantial amount of energy is converted from the barotropic to the baroclinic tide, as SHA shows. However, low-mode generation is an internal, adiabatic conversion, and, unlike parameterized drag, does not directly remove energy from the system as a whole. It is computationally infeasible for our hydrostatic, two-layer model to resolve the generation and breaking of internal waves at very small scales. Hence the need apparently remains for parameterized drag, which we have acting on the lower layer flow rather than the barotropic flow. Since the vertically integrated kinetic energy in the one- and two-layer simulations is the same to within 10 percent, the amount of parameterized drag required to obtain accurate tides is similar in the two cases.

5. Surface elevation accuracy of multi-constituent simulations

In the two subsections to follow we discuss the surface elevation accuracy of a 1/2° one-layer ten-constituent simulation and a 1/2° two-layer ten-constituent simulation. Both employ η_{MEM} to iterate the self-attraction and loading term. In both cases, sea-surface height discrepancies against the Ray results were very similar in the second and third iterations, and we take results from the third iteration. Our topographic drag scheme was used, augmented by a multiplicative factor of 7 in the one-layer case and 8 in the two-layer case.

5.1. One-layer multi-constituent run

The third column of Table 2 lists values of D in the one-layer run, measured against the Ray data in deep waters equatorward of 66°. The signal S for each constituent is given in the second column. The RSS anomaly of all the constituents is 9.54 cm, slightly lower than the 10.1 cm found by JS. Our M_2 and S_2 elevation discrepancies are slightly higher than those in JS. Numbers in parentheses denote percentage of sea-surface height variance captured. Although the discrepancies for the various constituents differ greatly in absolute terms, the percent variance captured is confined to a relatively narrow range (83.8–94.4) across the constituents. In contrast, the K_1, O_1, P_1, and Q_1 solutions in JS capture 67, 81, 80, and 25 percent of the variance, respectively. In terms of percent variance captured, our semi-diurnal constituents, especially S_2 and K_2, are less accurate than our diurnal tides. Our relatively poor S_2 results may be due to the air tide forcing at that frequency, which we have not accounted for (Rui Ponte and Gary Egbert, personal communication 2004). The percent variance captured is also relatively low for M_m and M_f. This is probably due to contamination by other low-frequency oceanic motions (Desai et al., 1997; Egbert and Ray, 2003b) in the Ray data, rather than to any inherent difficulty in simulating long-period tides. Our long-period elevations match those in the forward long-period simulations of Egbert and Ray (2003b) much more closely. The elevation

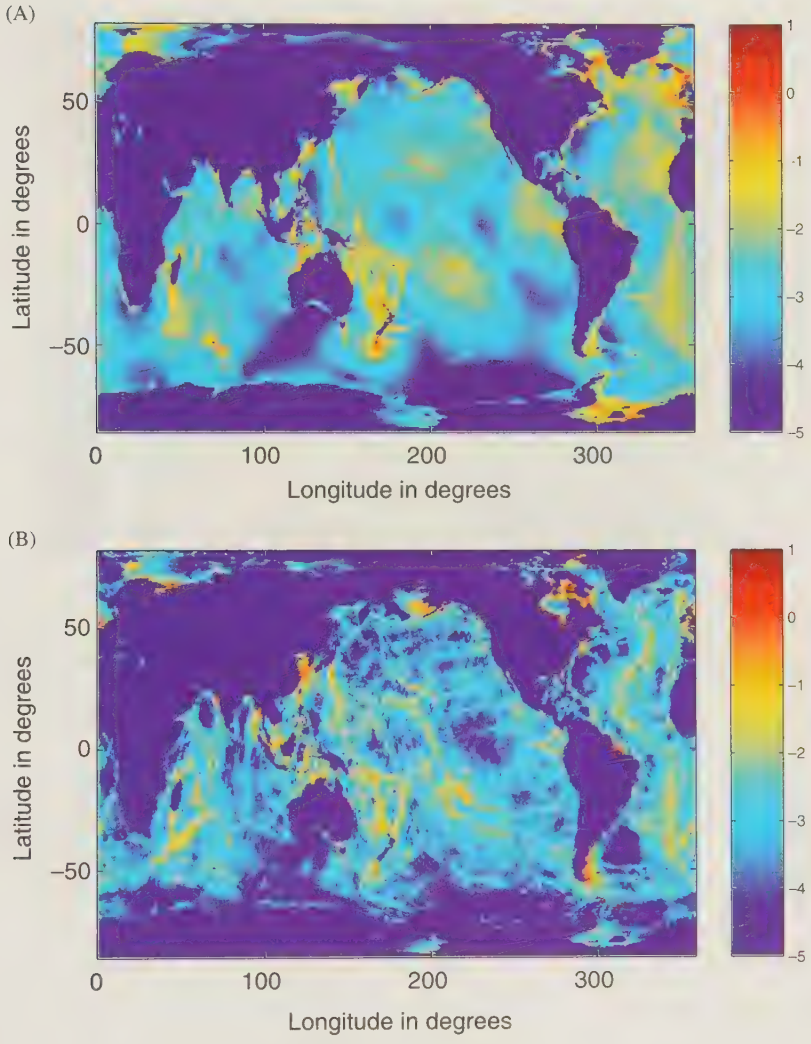


Fig. 5. Log_{10} of time-averaged parameterized dissipation (in W m^{-2}) in 1/2° one-layer M_2-only run with (A) optimally tuned c_d value of 0.64 and (B) optimally tuned topographic drag scheme (multiplicative factor of 7).

discrepancies of the four semi-diurnal and four diurnal constituents in our one-layer forward model are all slightly lower when measured against TPXO6.2 than when measured against GOT99.2, such that the RSS reduces from 9.54 to 9.42 cm.

Elevation discrepancies against the set of 102 pelagic tide gauges are shown in Table 3. For the sake of comparison, the (much lower) GOT99.2 discrepancies with the tide gauge data are also shown. The elevation errors of our one-layer forward model against tide gauges are larger than the errors against GOT99.2 for all constituents except K_2. However, across all constituents, the signal sampled by the gauges is larger than the globally averaged signal (compare the second columns of Tables 2 and 3), such that the percent of the RSS signal variance captured is slightly higher (91.9 versus 91.0 percent) when tide gauges are used as the benchmark. As in the comparisons with GOT99.2, our model captures diurnal elevations better than semidiurnal elevations, and again the poorest performance is in S_2 and K_2. The

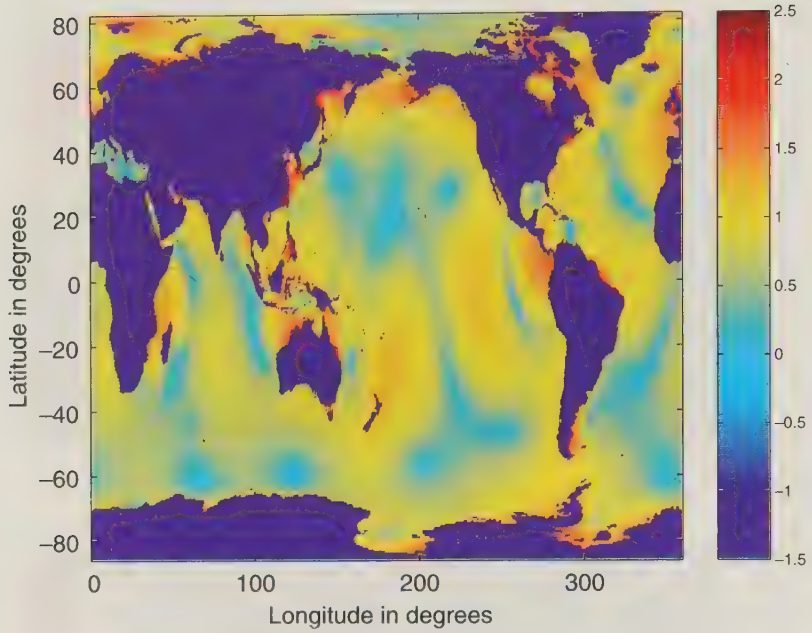


Fig. 6. Log_{10} of time-averaged sea level errors $\sqrt{\langle(\eta - \eta_{\mathrm{GOT99.2}})^2\rangle}$ (in cm) for 1/2° one-layer M_2-only experiment having topographic drag with multiplicative factor of 7.

Table 2
Second column: time- and area-averaged sea-surface height signals S calculated from GOT99.2 (for the diurnals and semi-diurnals) and from empirical mappings (for M_f and M_m). Third and fourth columns: Time- and area-averaged sea-surface height discrepancies D of 1/2° multi-constituent one- and two-layer forward simulations with respect to GOT99.2 (for the diurnals and semi-diurnals) and to the empirical mappings (for M_f and M_m).

Constituent	Signal S (cm)	One-layer D (cm)	Two-layer D (cm)
M_m	0.49	0.17 (87.3)	0.17 (87.3)
M_f	0.90	0.27 (91.0)	0.28 (90.0)
Q_1	1.39	0.36 (93.2)	0.38 (92.7)
O_1	6.61	1.57 (94.4)	1.53 (94.7)
P_1	3.13	0.77 (93.9)	0.64 (95.9)
K_1	9.54	2.45 (93.4)	1.88 (96.1)
N_2	5.65	1.51 (92.9)	1.37 (94.1)
M_2	26.69	7.76 (91.6)	7.26 (92.6)
S_2	10.57	4.26 (83.8)	4.12 (84.8)
K_2	2.97	1.08 (86.8)	1.05 (87.6)
RSS	31.82	9.54 (91.0)	8.90 (92.2)

All quantities are computed in waters deeper than 1000 m and equatorward of 66°. Numbers in parentheses denote percentage of sea-surface height variance captured in the forward simulations.

fourth column of Table 4 lists the elevation discrepancies of our one-layer forward model against TPXO6.2 in waters deeper than 1000 m and latitudes south of 66°S. For comparison, the TPXO6.2 signals, and the discrepancies between GOT99.2 and TPXO6.2, are also listed. In this comparison, once again we fare better with diurnal tides than with semidiurnal tides. However, the

Table 3
Time and station-averaged sea-surface height signals S at the set of 102 pelagic tide gauges used in Shum et al. (1997), and sea-surface height discrepancies D of GOT99.2 and of our 1/2° multi-constituent one- and two-layer forward simulations with respect to the gauges

Constituent	Signal S (cm)	GOT99.2 D (cm)	One-layer D (cm)	Two-layer D (cm)
Q_1	1.62	0.28	0.43 (92.8)	0.40 (93.9)
O_1	7.76	0.89	1.79 (94.7)	1.62 (95.7)
P_1	3.62	0.37	0.84 (94.6)	0.73 (95.9)
K_1	11.26	1.02	2.90 (93.3)	2.26 (96.0)
N_2	6.86	0.65	1.95 (91.9)	1.83 (92.9)
M_2	33.22	1.48	9.33 (92.1)	8.75 (93.1)
S_2	12.62	1.03	4.44 (87.6)	4.25 (88.7)
K_2	3.43	0.43	1.00 (91.4)	0.96 (92.2)
RSS	39.04	2.43	11.14 (91.9)	10.36 (93.0)

Numbers in parentheses denote percentage of sea-surface height variance captured in the forward simulations. Our calculated GOT99.2 discrepancies agree with those in Table 2 of Ray (1999), except that he applied an air tide correction to the bottom pressure recorder S_2 data which reduces the S_2 discrepancy to 0.94 cm. The air tide correction is not present in the Shum et al. dataset we used.

Table 4
Southern ocean time- and area-averaged sea-surface height signals S calculated from TPXO6.2 and sea-surface height discrepancies D of GOT99.2 and of our 1/2° multi-constituent one- and two-layer forward simulations with respect to TPXO6.2

Constituent	Signal S (cm)	GOT99.2 D (cm)	One-layer D (cm)	Two-layer D (cm)
Q_1	3.79	0.40	0.50 (98.2)	0.41 (98.8)
O_1	17.60	3.89	2.11 (98.6)	1.72 (99.0)
P_1	5.73	0.90	1.52 (93.0)	0.78 (98.1)
K_1	18.11	2.66	5.06 (92.2)	2.53 (98.1)
N_2	3.63	0.94	1.52 (82.4)	1.36 (86.0)
M_2	18.96	2.01	6.75 (87.3)	5.99 (90.0)
S_2	13.35	1.04	5.09 (85.5)	5.05 (85.7)
K_2	3.79	0.41	1.36 (87.1)	1.32 (88.0)
RSS	35.36	5.41	10.40 (91.4)	8.66 (94.0)

All quantities are computed in waters deeper than 1000 m and south of 66°S. Numbers in parentheses denote percentage of sea-surface height variance captured in the forward simulations.

difference is more striking in the Antarctic. The semi-diurnal constituents are never captured at better than 87.3 percent, while the diurnal constituents are all captured at 92.2 percent or better.

Table 5 lists the global tidal working rates computed from T/P-constrained models by Egbert and Ray (2003a) alongside the working rates, computed from (1) over all model gridpoints (deep and shallow), in our one-layer multi-constituent run. With the exception of the Q_1 rate, our one-layer forward model working rates match those in the T/P-constrained models to within 8 percent or better. The largest discrepancy is in the M_2 rate, which is 0.2 TW too large, as is the total rate summed over all constituents. A rough estimate of the parameterized dissipation in the abyss of our multi-constituent run can be made by assuming that the topographic drag decay rate is spatially uniform, so that the last term in (3) is $-r\vec{u}$ rather than $\overline{T}\vec{u}/\rho_0(H+\eta)$. Multiplication of $-r\vec{u}$ by $\rho_0(H+\eta)\vec{u}$ and global integration would yield $-\rho_0 r \iint (H+\eta)\vec{u}\cdot\vec{u}\,dA = -2r\,\mathrm{KE}$ for the decay rate of tidal energy. The spatially-averaged decay rate computed from (19) in waters deeper than 1000 m when the optimal multiplicative factor of 7 is used is $(2.0\ \mathrm{days})^{-1}$, very close to the

Table 5
The global working rates of the tidal constituents

Constituent	Egbert and Ray global work (TW)	One-layer global work (TW)	Two-layer global work (TW)
Q_1	0.007	0.008 (+14)	0.009 (+29)
O_1	0.173	0.181 (+5)	0.198 (+14)
P_1	0.035	0.034 (−3)	0.034 (−3)
K_1	0.343	0.329 (−4)	0.334 (−3)
N_2	0.110	0.115 (+5)	0.114 (+4)
M_2	2.435	2.632 (+8)	2.577 (+6)
S_2	0.376	0.392 (+4)	0.380 (+1)
K_2	0.030	0.029 (−3)	0.028 (−7)
Total	3.508	3.719 (+6)	3.672 (+5)

The second column is from Table 1 of Egbert and Ray (2003a), computed from data-constrained solutions. The third and fourth columns are the workings in our 1/2° one- and two-layer multi-constituent forward simulations. Numbers in parentheses indicate the percentage differences of the model results with respect to the Egbert and Ray results.

$(2.1 \text{ days})^{-1}$ rate computed from (20) in the optimal run in JS. If one multiplies $(2.0 \text{ days})^{-1}$ by twice the abyssal model kinetic energy in our optimally tuned one-layer multi-constituent experiment, one estimates 2.3 TW for the parameterized abyssal dissipation. This is slightly larger than the 1.85 TW abyssal dissipation we compute from an energy budget analysis (see SHA for details) of our one-layer multi-constituent run. The energy budget analysis provides working and dissipation rates at every model gridpoint, as the model is running, while (1) holds only when integrated globally. The energy budget analysis thus yields a working rate that is independent from that obtained by summing the constituent values obtained via (1). The two estimates of the global total working rates agree with each other and with the global total dissipation estimate to better than 0.01 TW. The energy budget is computed through multiplication of the momentum equation by $(H + \eta)\vec{u}$ as the model is running, and hence does not separate out the constituents. The abyssal dissipations in our one-layer model are significantly larger than those found by Egbert and Ray (2000, 2001, 2003a) in T/P-constrained solutions, and by the JS and ERB one-layer forward models.

5.2. *Two-layer multi-constituent run*

The fourth column of Table 2 lists the elevation discrepancies against the Ray results in our two-layer multi-constituent run. The discrepancies are lower than those in the one-layer multi-constituent run across all constituents save the three smallest. The RSS error over all constituents is reduced from 9.54 to 8.90 cm (or, from 9.42 to 8.78 cm when measured against TPXO6.2). Again, the semi-diurnal tides, especially S_2 and K_2, are less accurate (in terms of percentage of sea-surface height variance captured) than the diurnal tides. Table 3 shows that the improvement in elevation accuracies in the two-layer forward model over the one-layer forward model also holds when pelagic tide gauges are used as the benchmark. The addition of another layer yields an extra percent of sea surface height variance captured in the RSS signal, just as it does when elevations are measured against GOT99.2. Table 4 shows that the inclusion of another layer also lowers elevation errors across all constituents when TPXO6.2 is used as a benchmark in waters deeper than 1000 m and south of 66°S. The increase in percent variance captured is greatest for P_1 and K_1, as it was when measuring the one and two layer models against the other two benchmarks. The diurnal elevations in the two-layer forward model match the TPXO6.2 elevations in the open Antarctic at the level of 98 to 99 percent variance captured. The P_1 and K_1 elevations in GOT99.2 match those in TPXO6.2 slightly less well in the Antarctic than do the P_1 and K_1 elevations in our two-layer forward simulations, and our two-layer model apparently

captures the O_1 elevations significantly better. The GOT99.2 semi-diurnal elevations match TPXO6.2 much more closely than do the semi-diurnal elevations in our forward model. With the notable exception of the O_1 global working rate, the constituent global working rates of the two-layer solution are generally slightly more accurate than those in the one-layer solution, and the working rate over all constituents is slightly closer to the Egbert and Ray numbers (Table 5). Note that although we use the energy budget analysis of SHA, we do not employ mode-splitting, and consequently our energy inputs and outputs in the two-layer case balance much more closely (to 0.01 TW) than those in SHA.

Fig. 7 displays a map of the elevation discrepancies between our M_2 and K_2 solutions in the two-layer multi-constituent run and TPXO6.2. In both cases, the elevation difference is scaled by the globally averaged signal S of TPXO6.2, computed

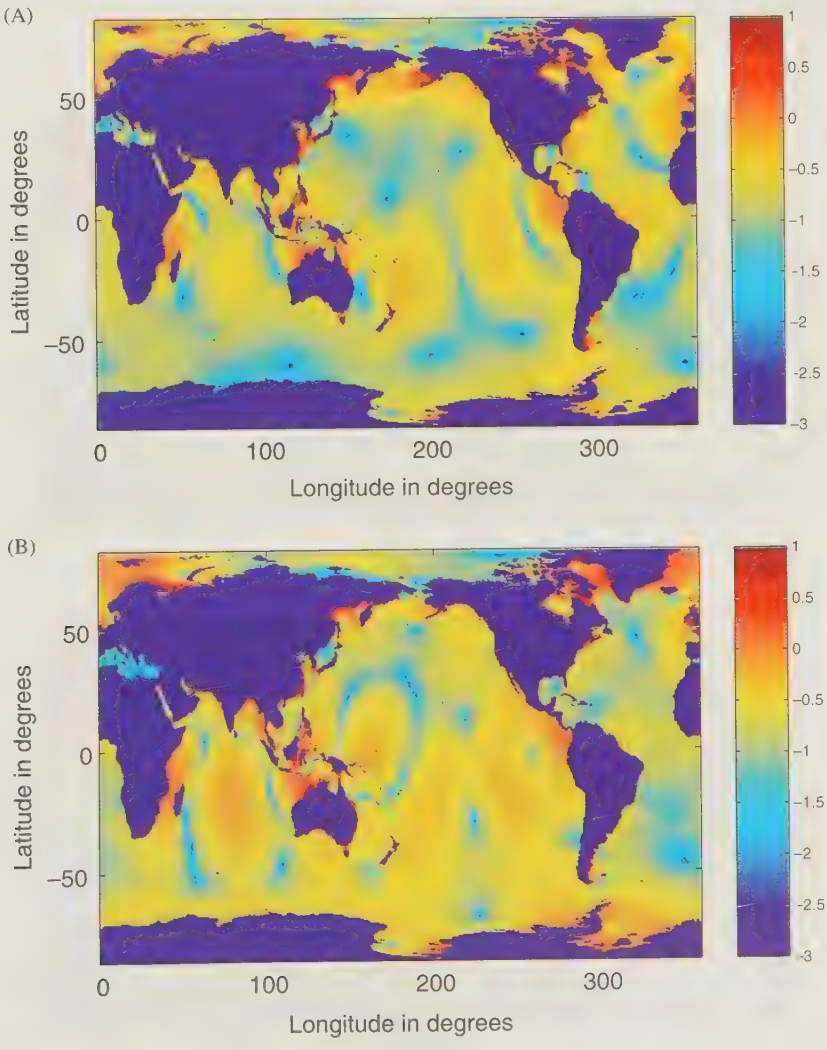


Fig. 7. (A) Log_{10} of normalized M_2 sea-level errors $\sqrt{\langle(\eta-\eta_{\mathrm{TPXO6.2}})^2\rangle}/S$ in the 1 2 two-layer multi-constituent run, where brackets denote time-averaging, and S is the time- and area-averaged M_2 elevation signal of TPXO6.2 in waters deeper than 1000 m and equatorward of 66°. (B) Same but for K_2.

in waters deeper than 1000 m and equatorward of 66°. When unscaled (not shown) the M_2 map closely resembles the M_2 error map in the one-layer simulations, and is virtually indistinguishable, except at high latitudes, from an error map computed with respect to GOT99.2. The locations of large errors in the M_2 and K_2 maps are quite different. A frequency sensitivity in the appearance of error maps is also seen in Fig. 8, which plots the scaled elevation differences in the two-layer K_1 and O_1 solutions. In the diurnal maps "dimples"—the surface elevation signature of first mode baroclinic tides—can be seen between 30°S and 30°N, the latitudes of active diurnal internal tides. Correlations between the error maps for the various constituents were calculated. The highest correlation coefficients are between the S_2 and K_2 error maps (0.98), and the K_1 and P_1 error maps (0.97), consistent with the fact that these pairs lie close together in frequency. Coefficients for other diurnal pairs and semidiurnal pairs lie between 0.80 and 0.94. The frequency sensitivity of the error maps is reminiscent of the frequency sensitivity seen in the normal mode maps of Platzman et al. (1981). Correlation coefficients between maps of constituent errors and the

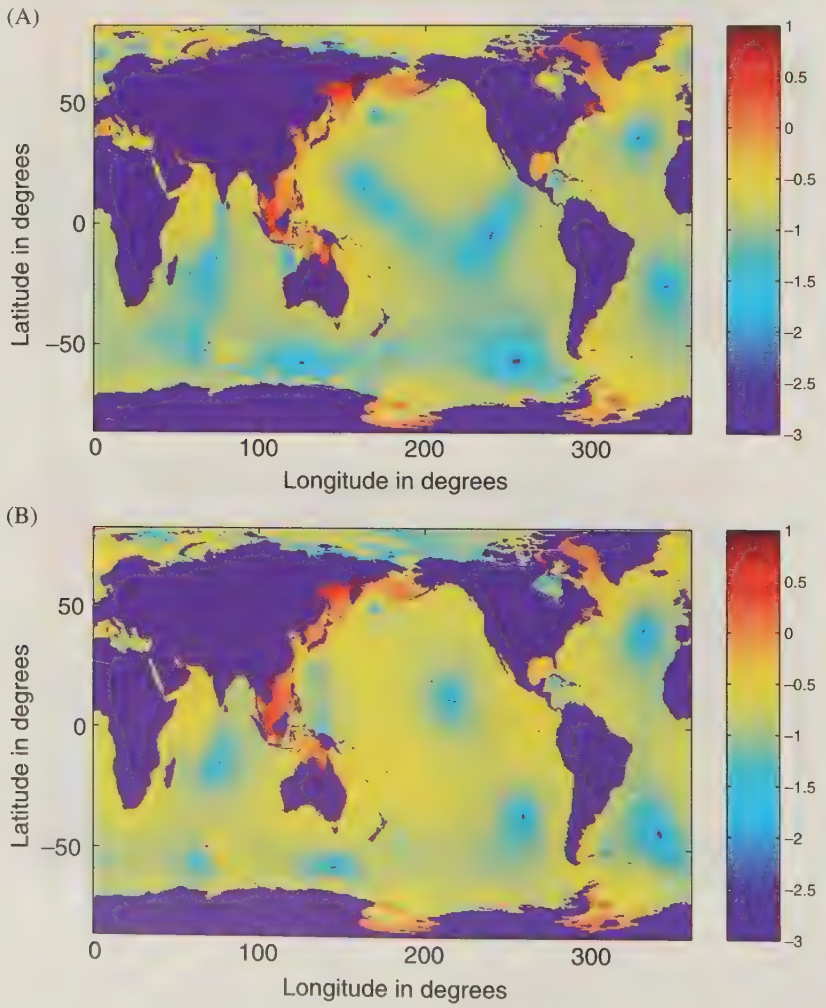


Fig. 8. As in the previous figure but for (A) K_1 and (B) O_1.

corresponding constituent signals ranged from 0.76 to 0.88, indicating that large errors often occur where signals are large. Comparison of Fig. 8 with Fig. 9, which shows the K_1 and O_1 error maps in the one-layer multi-constituent run, indicates that baroclinicity alters the solution globally. Thus, the increase in accuracy seen in our two-layer K_1 solution in the Antarctic, which takes place despite the fact that our chosen g' values are more appropriate for mid-latitudes, may be due more to the greater accuracy of the solution elsewhere (which serves as a boundary forcing for southern latitudes) than to local effects.

Fig. 10 shows the amplitude and phase of K_1 surface elevations in our one-layer multi-constituent solution, in our two-layer multi-constituent solution, and in GOT99.2, along longitudes 140°E in the western Pacific and 330°E in the Atlantic, where relatively large reductions in the surface elevation errors are realized with baroclinicity (compare Figs. 8A and 9A). In the two-layer solution, the direct signature of the baroclinic tide is evident as small-amplitude, small-scale oscillations superimposed on the large-scale surface elevation fields. The inclusion of baroclinicity shifts the large-scale surface elevation field

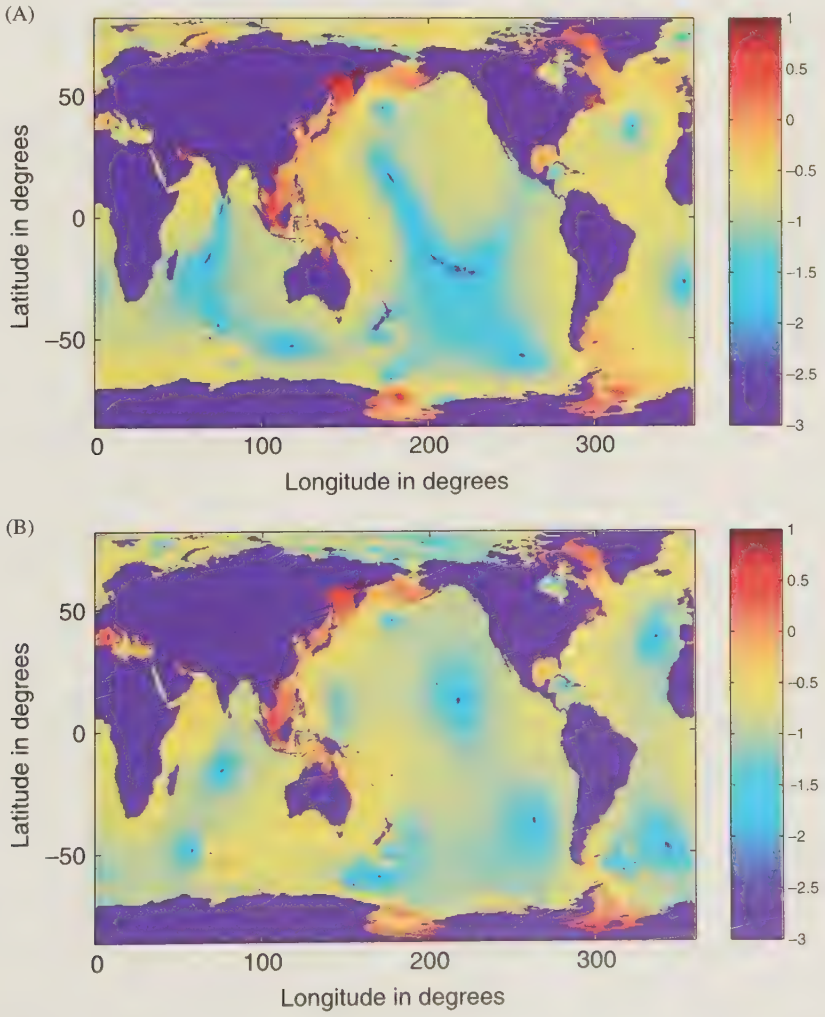


Fig. 9. As in the previous two figures but for (A) K_1 and (B) O_1 solutions in the 1/2° one-layer multi-constituent run.

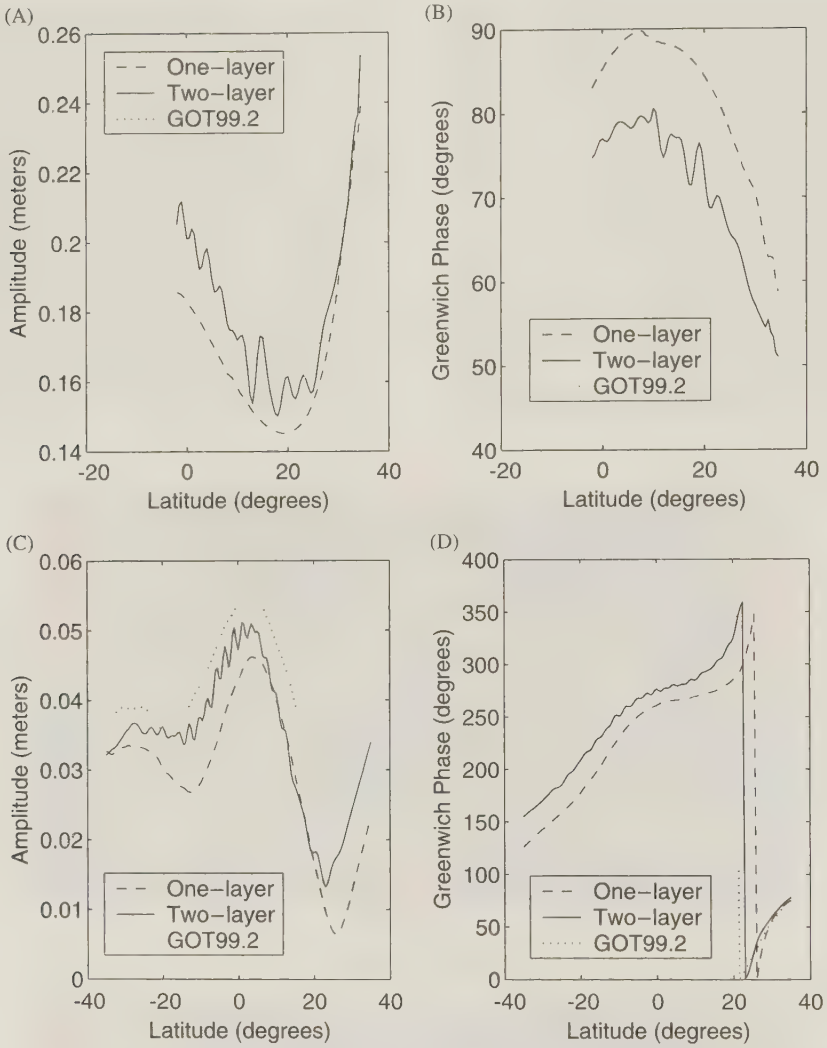


Fig. 10. Amplitude along (A) 140°E and (C) 330°E, and Greenwich phase along (B) 140° E and (D) 330° E of K_1 solutions.

towards the GOT99.2 elevations. It is this shift in the barotropic tide induced by baroclinicity, rather than direct surface effects, that lies behind the lower elevation errors in the two-layer run. Note that the baroclinic signature in the surface elevations is not evident in GOT99.2. As explained in Ray and Mitchum (1996, 1997), most T/P-constrained tide models filter out the baroclinic tide by averaging across tracks. Fig. 10 suggests that small spatial scales should perhaps be filtered out of the surface elevations in our two-layer model before comparison is made to elevations in T/P-constrained models. However, we have chosen for simplicity not to perform such filtering, and we note that this choice is appropriate for comparison to pelagic tide gauge data, which does not filter out the surface signature of baroclinic tides.

6. Interfacial displacements and velocities in two-layer M_2 run

Fig. 11 displays a snapshot of the interfacial height in a 1/4° two-layer M_2-only experiment with

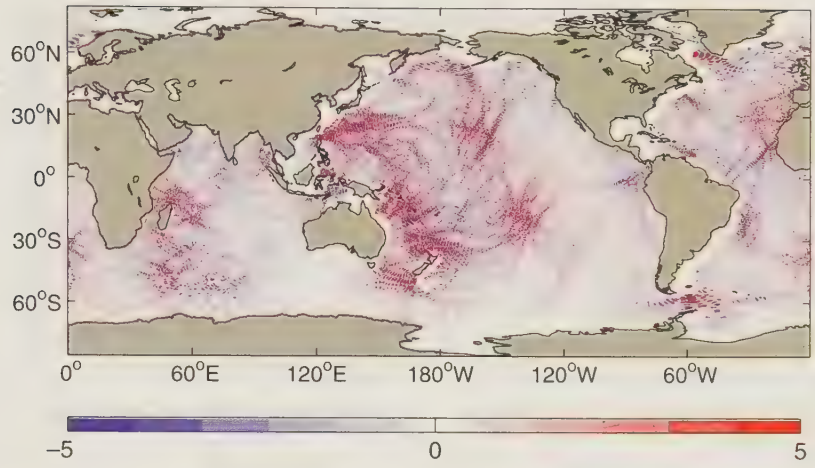


Fig. 11. Snapshot of interfacial height (m) in 1/4° M_2-only two-layer run with a multiplicative factor of 8, taken after the model has equilibrated. The lower layer only exists where the water column depth exceeds 700-η_2 m.

a multiplicative factor of 8. The experiment has been iterated with respect to η_{SAL}. The snapshot was taken long after model equilibration was achieved. "Hotspots" for the generation of low-mode baroclinic tides, such as Hawaii, show up clearly. The globally integrated interfacial APE $= (1/2)\rho_0 \iint g' \eta_2^2 \, dA$ is 2.3 times lower than that of the two-layer nominal c_d-only experiment (i.e., the run without any topographic drag). The interfacial displacements in Fig. 8 of SHA are larger than those in our Fig. 11 (note the difference in scales), consistent with the lack of topographic drag in SHA. Topography was mapped onto the coarse model grids in different ways in the two papers, and that also may affect the interfacial displacements. Tidal beams do not appear to propagate as far from their sources in experiments with topographic drag as they do in experiments without topographic drag (again, compare Fig. 11 with Fig. 8 in SHA). Fig. 12 plots the spatial pattern of the ratio of modal kinetic energy densities $\langle (H_1 + H_2 + \eta_1)|\vec{u}_{BC}|^2 \rangle / \langle (H_1 + H_2 + \eta_1)|\vec{u}_{BT}|^2 \rangle$ in the 1/4° M_2-only run of Fig. 11. As in the interfacial displacement field, fine spatial structure is evident, with local values of the modal kinetic energy density ratio often being as large as 3. The ratio of temporally averaged global baroclinic to barotropic kinetic energy $\langle KE_{BC} \rangle / \langle KE_{BT} \rangle$ is 0.13; however, this ratio is sensitive to the amount of horizontal friction present. A simulation with $K_H = 250 \, m^2 s^{-1}$ yielded a baroclinicity ratio of 0.20, while surface elevations were virtually unaffected. Fig. 13 plots the ratio of temporally averaged upper and lower layer squared speeds $\langle |\vec{u}_1|^2 \rangle / \langle |\vec{u}_2|^2 \rangle$ in the same 1/4° M_2-only run. Once again, much fine structure is evident, and local values of the ratio are often as high as ten. The larger tidal currents in the upper layer are likely due to the action of friction on the bottom layer as well as to the surface intensification of the baroclinic mode in the presence of thin thermoclines. Arbic and Flierl (2004) explores similar effects in idealized models of the mesoscale eddy field.

7. Summary and discussion

Over the last two decades, open-ocean tidal elevations have been mapped with high accuracy in models constrained by either tide-gauge or altimetric data. Some of these models are purely empirical, while others have hydrodynamical cores and assimilate data. Models of the latter type, for instance, Le Provost and Lyard (1997), have been able to accurately map tidal elevations without including substantial parameterized drag in the open-ocean parts of their domain. However, both

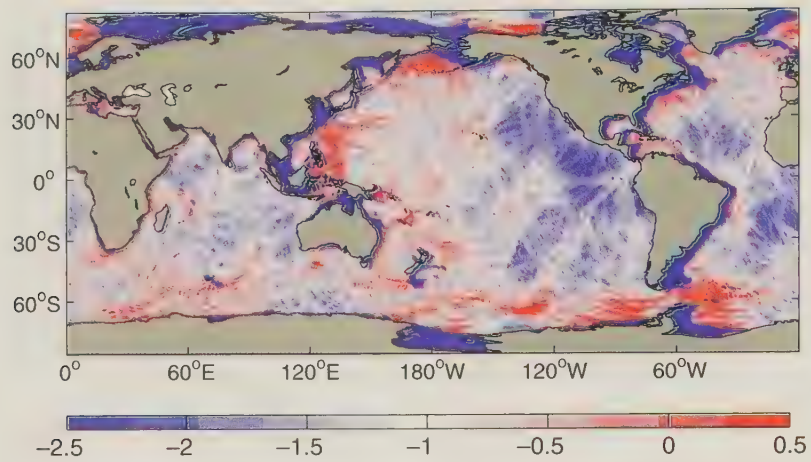


Fig. 12. Log_{10} of $\langle (H_1 + H_2 + \eta_1)|\vec{u}_{\mathrm{BC}}|^2\rangle / \langle (H_1 + H_2 + \eta_1)|\vec{u}_{\mathrm{BT}}|^2\rangle$, where brackets denote time-averaging, in the M_2 simulation of the previous figure. Areas with resting water column depths less than 700 m are shaded dark blue.

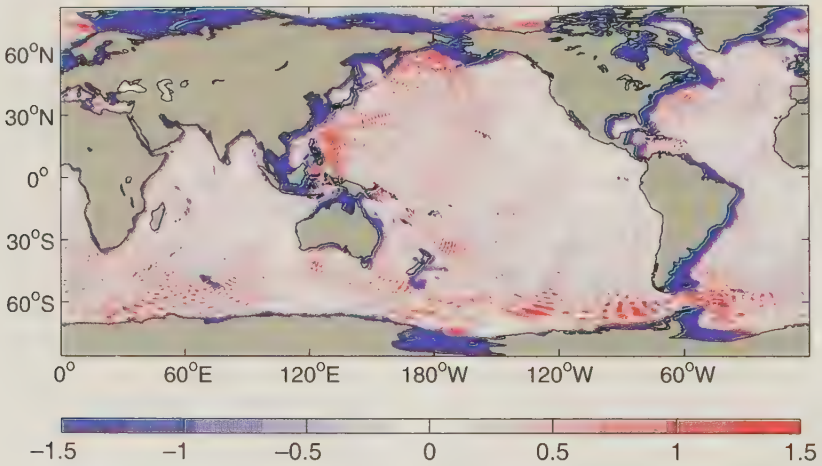


Fig. 13. Log_{10} of $\langle |\vec{u}_1|^2\rangle / \langle |\vec{u}_2|^2\rangle$, where brackets denote time averaging, in the M_2 simulation of the previous two figures. Areas with resting water column depths less than 700 m are shaded dark blue.

in situ observations and inferences from other T/P-constrained models suggest that substantial tidal dissipation (about 1 TW) takes place in the abyss. JS, CL, ERB, and the present paper have all found that the elevation accuracy of forward modeled tidal elevations is improved when substantial parameterized drag is included in the abyss. In this paper we have put forth a simple explanation for the sensitivity of model elevation accuracy to the parameterized drag. Both the APE and KE of semidiurnal tides depend on the amount of abyssal drag. (The KE of both diurnal and long-period tides is sensitive to drag. The APE of diurnal tides is less sensitive to drag than is the APE of semidiurnal tides, and APE of long-period tides shows little sensitivity). Elevation errors are minimized when the model APE is near the observed value. Fairly accurate tidal elevations can be obtained in runs without topographic drag, in which other frictional parameters such as c_d or

K_H are treated as tunable, but only if c_d or K_H are made far larger than values that are physically plausible. Using topographic drag is preferable because it is physically motivated and concentrates the parameterized dissipation over areas of rough topography, as is seen in observations. Moreover, large K_H values would prohibit modeling of features which have small horizontal scales, such as coastal tides, baroclinic tides, and mesoscale eddies. Usage of topographic drag may open up the prospect for development of global models that capture both tides and the slowly varying oceanic general circulation with some degree of accuracy.

JS used the topographic drag formula $\rho_0 N k h^2 \vec{u}/2$ for steady flow over monochromatic topography $h \sin(kx)$. They had no theoretical guidance behind their specific choices of h and k in the face of real topography, which has a spectrum of topographic heights and wavenumbers. Our scheme begins with an exact linear analysis for the drag on an oscillatory background flow over arbitrary small-amplitude topography. The linear analysis is supplemented by dimensional reasoning, based on arguments in the atmospheric literature, which differentiates between linear drag arising from radiating waves and quadratic drag arising from turbulence that results when flow deflects around topography. We have found that our calculated drag must be supplemented with a multiplicative factor, which we tune to minimize anomalies of our forward-modeled elevations against observations. Our current interpretation is that the multiplicative factor accounts for the unresolved small scales in the input topographic datasets, on which both flow-splitting and generation of breaking high mode internal waves likely take place. We find an optimal multiplicative factor of 7 in the one-layer M_2-only case, and then use this factor across all constituents in one-layer multi-constituent simulations. It is the only free parameter in our model. JS also had a tunable parameter (topographic wavenumber k). The area-averaged decay rate in waters deeper than 1000 m lies between $(2.0 \text{ days})^{-1}$ and $(2.2 \text{ days})^{-1}$ in our optimally tuned run, the optimally tuned run in JS, and the optimally tuned run in Hirose et al. (2001), who examined the accuracy of models of wind-driven barotropic motions. The consistency in decay rates suggests that topographic drag probably affects both wind-driven and tidal motions. Arbic and Flierl (2004) find that eddy vertical and horizontal structure in idealized models of baroclinically unstable geostrophic turbulence matches that of mid-ocean mesoscale eddies best when bottom friction is fairly strong. The current paper and others have argued that topographic drag may be a dominant source of this bottom friction.

Consistent with ERB, we have documented the lack of convergence in straightforward iterations of the self-attraction and loading term, and demonstrated that convergence can be achieved when a memory of the tidal elevations in prior iterations is retained. We have therefore obtained self-consistent forward solutions for the global tide. In waters deeper than 1000 m and in latitudes covered by the T/P altimeter, the sea-surface height discrepancy D against GOT99.2 in our optimally tuned ten-constituent one-layer forward model is 9.54 cm, slightly lower than the 10.1 cm found in JS. Across different constituents, the absolute values of D vary widely, but the percent variance captured is consistently high (84 percent or greater). The discrepancies of our one-layer multi-constituent forward model against the set of 102 pelagic tide gauges used in Shum et al. (1997) is 11.14 cm, substantially larger than the error measured against GOT99.2. However, the tidal signals sampled by the gauges are larger than the globally averaged signals in altimeter-constrained models, such that the percent elevation variance captured by our one-layer forward model is similar whether tide gauges or T/P-constrained models are used as the benchmark.

As far as we know, this paper and SHA are the first to publish simulations of baroclinic tides on a global scale. Both papers display global snapshots of the interfacial displacements, which show a great deal of fine structure including tidal beams. Both papers also document a substantial impact of baroclinicity on the tidal velocity field. Here we have mapped the ratio of temporally averaged baroclinic to barotropic kinetic energy densities, local values of which are as high as 3, and the ratio of temporally averaged squared upper to lower

layer current speeds, local values of which are as high as 10. SHA calculate conversion rates of barotropic to baroclinic tidal energy. In the present paper we have shown that the addition of another layer improves the surface elevation accuracy of our forward model across all of the larger constituents, when each of three different benchmarks is used. The improvement in surface elevation accuracy is not due to the direct elevation signature of the baroclinic tide, which is small, but to shifts in the surface elevations induced by baroclinicity. The RSS sea-surface height discrepancy of our multi-constituent forward tide model against GOT99.2 reduces from 9.54 to 8.90 cm, and the discrepancy against pelagic tide gauges reduces from 11.14 to 10.36 cm. Improvements in elevation accuracy from the one-layer to two layer model are also seen in waters deeper than 1000 m and south of 66°S, where we consider the TPXO6.2 model to be our benchmark. Against all three benchmarks, especially the Antarctic one, the improvement in elevation accuracy with baroclinicity is most noticeable for the P_1 and K_1 constituents. Our two-layer model apparently captures 98 to 99 percent of the open southern ocean diurnal height variance. The multiplicative factor for topographic drag was re-tuned in the two-layer case and the optimal value is 8. This is nearly the same as the optimal factor in the one-layer case, against our initial expectations that the explicitly resolved generation of low modes would allow less parameterized drag. In a follow-up paper we will combine the low-mode conversion diagnostic of SHA with the topographic drag used in the current paper to more thoroughly explore the roles of parameterized drag and resolved low-mode conversion in baroclinic tide models. We will explore the possibility that direct removal of energy from the low-mode baroclinic tides via parametric subharmonic instability (MacKinnon and Winters, submitted 2004) or other mechanisms, might allow a reduction in the amount of parameterized topographic drag needed at the bottom.

Maps of elevation error reveal large regions over which our forward model differs substantially from data-constrained models. We believe therefore that our forward model is still missing some important physics. It is possible that the lack of other oceanic motions in our model degrades its accuracy. This, however, seems unlikely to us since the largest elevation errors do not generally occur in regions of strong currents. Furthermore, as measured by percentage of sea-surface height variance captured, our forward-modeled diurnal tides are more accurate than our semi-diurnal tides, despite the fact that diurnal frequencies lie closer to those of other energetic oceanic currents. Desai et al. (1997) found that the remaining errors in empirical ocean tide models are due largely to the oceanic general circulation. However, these errors are much smaller than those in our forward model. The general circulation affects the global distribution of density, which in turn affects the internal tides. Here we have used a single value of reduced gravity, and a single value of the resting interfacial depth, to characterize stratification, but stratification in the actual ocean is more complicated. Usage of the simpler topographic drag scheme of JS, or of large c_d or K_H values, produces elevation accuracies comparable to those achieved with our topographic drag scheme. ERB tried three different topographic drag schemes in their forward model and found little effect on elevation accuracy. We therefore doubt that changes to the drag scheme will make a practical difference in the elevation accuracy. Improvements in the underlying topographic datasets might reduce errors in forward models. The globally averaged sea-surface height discrepancy in preliminary M_2 experiments with the IOC et al. (2003, "GEBCO") topographic dataset is larger than in experiments using the Coward dataset, which was constructed primarily from Smith and Sandwell (1997). Model sensitivity to the underlying bathymetry, along with the related sensitivities to grid resolution and to the types and widths of filters used to map topographic datasets onto model grids, will be explored in a future paper. Another potential improvement to our forward tide model would be to use load calculations from a solid-earth model that incorporates lateral inhomogeneities (Tom Jordan, Tony Dahlen, Jerry Mitrovica, and Konstantin Latychev, personal communication 2003).

Acknowledgements

BKA thanks Laurie Padman and an anonymous reviewer for very thorough and helpful reviews; Richard Ray, Anand Gnanadesikan, and Steve Griffies for also improving the manuscript with their comments; Chris Garrett for steering him towards tides as an interesting problem; Gary Egbert and Richard Ray for patiently answering many questions about tides and for providing several useful datasets; Steve Jayne for much helpful practical advice; Raffaele Ferrari, Myrl Hendershott, Larry Horowitz, Isaac Held, Chris Kerr, Jody Klymak, Jennifer MacKinnon, Jim McWilliams, Jonathan Nash, David Nolan, Jonas Nycander, Isidoro Orlanski, Rob Pinkel, Kurt Polzin, David Sandwell, Walter H.F. Smith, John Toole, Andrew Wittenberg, Carl Wunsch, and Bill Young for useful discussions; Jeff Durachta, Brian Gross, Amy Langenhorst, John Sheldon, Buz Taylor, Hans Vahlenkamp, and Tim Yeager for help with computer software and hardware issues; and the Princeton University/GFDL Visiting Scientist Program for financial support during the course of this research. Simulations were performed on the supercomputer cluster at GFDL.

Appendix. Details of topographic drag scheme

We begin with the linear analysis for oscillatory background flow, and then discuss the augmentations from the dimensional analysis for drag due to low-level breaking. We assume for simplicity that the background flow is characterized by a single frequency ω. For simplicity, in this paper we take ω to be that of M_2 in all of our simulations, including those with multiple constituents. For the linear analysis, we assume that $|\vec{u}||\vec{k}| \ll \omega$, also for simplicity. (The linear solution in the general case is given by Bell, 1975a). Let the Fourier amplitudes of the topography be $\widehat{h}(\vec{k})$, where $\vec{k} = (k, l)$ is the horizontal wavenumber. Let $u_{\vec{k}} = \vec{u} \cdot \vec{k}/|\vec{k}|$ be the component of the background flow $\vec{u}$ along the gradient of the terrain component having wavenumber $\vec{k}$. The linear boundary condition is

$$\widehat{W}(\vec{k}) = iu_{\vec{k}}|\vec{k}|\widehat{h}(\vec{k}), \tag{24}$$

where $\widehat{W}(\vec{k})$ is the Fourier transform of $W(x, y) = w(x, y, z = 0)$, the vertical velocity component at the bottom. We assume gradual vertical variations such that $\mathrm{d}(|m|^{-1})/\mathrm{d}z \ll 1$, where m is the vertical wavenumber, and also assume $m \gg |\mathrm{d}\overline{\rho}/\mathrm{d}z|/\overline{\rho}$, where $\overline{\rho}(z)$ is the vertical profile of density. By ignoring horizontal variations of the Brunt–Vaisala frequency N, neglecting vertical variations of $\overline{\rho}$, and considering hydrostatic scales $|\vec{k}| \ll m$, we can write the WKB approximation

$$\widehat{w}(\vec{k}, z) = \widehat{W}(\vec{k}) \exp\left(i \int_0^z m \, \mathrm{d}z'\right), \tag{25}$$

(e.g., Chapter 9 of Holton, 1972). For the oscillatory case

$$m = -N|\vec{k}|/\sqrt{\omega^2 - f^2}. \tag{26}$$

We have excluded nonhydrostatic waves by assuming that $\omega \ll N$. Note that throughout the analysis N refers to the value of N at the bottom, which we compute at every model gridpoint from Levitus et al. (1998).

From (24) and (25) it follows that, at the bottom,

$$\frac{\partial \widehat{w}}{\partial z} = \left(\frac{N|\vec{k}|}{\sqrt{\omega^2 - f^2}} \widehat{h\vec{k}}\right) \cdot \vec{u}. \tag{27}$$

Conservation of mass implies that $\partial w/\partial z = -\nabla \cdot \vec{u}'$, where $\vec{u}'$ is a perturbation velocity, and since the momentum equation implies

$$\left(\frac{\partial^2}{\partial t^2} + f^2\right)(-\nabla \cdot \vec{u}') = \overline{\rho}^{-1} \frac{\partial}{\partial t} \nabla^2 p', \tag{28}$$

where p' is a perturbation pressure, we may operate on (27) with the transform equivalent of $\overline{\rho}\nabla^{-2}(\partial^2/\partial t^2 + f^2)$ to obtain $\widehat{p} = \overline{\rho}(N_* \widehat{h}|\vec{k}|^{-1} i\vec{k}) \cdot \vec{u}$, where $N_* = N\sqrt{\omega^2 - f^2}/\omega$. If $\omega > |f|$, the Fourier synthesis is

$$p' = -\overline{\rho}\nabla\chi \cdot \vec{u}, \tag{29}$$

where χ is defined by

$$\chi(x, y) = -N_* \int\int \frac{\widehat{h}(\vec{k})}{|\vec{k}|} \exp(i\vec{k} \cdot \vec{x}) \, \mathrm{d}k \, \mathrm{d}l, \tag{30}$$

with $\vec{x} = (x, y)$. The spatial-transform equivalent is

$$\chi(x, y) = -\frac{N_*}{2\pi} \int\int \frac{h(\vec{x'})}{|\vec{x} - \vec{x'}|} \, dx' \, dy', \tag{31}$$

which is a slightly smoothed version of the terrain $h(\vec{x})$. The input topography must be filtered to retain only the topographic scales that force internal gravity waves.

The drag is determined by the pressure at the bottom according to

$$\tau = -p' \nabla h = [\overline{\rho} \nabla \chi (\nabla h)^{\mathrm{T}}] \vec{u}, \tag{32}$$

where the T superscript denotes the transpose operator. Formula (32), when multiplied by $\vec{u}$ and averaged over one tidal cycle, is equivalent to the modal energy conversion formulae in Nycander (2004, submitted), except that Nycander uses a filter to limit the length scales over which (31) is computed, instead of imposing a sharp cutoff at 40 km as we did. The factor in brackets, which we denote by $\overline{T}$, is a tensor that depends only on the amplitude, variance, and anisotropy of the topography. In the Appendix we let angle brackets denote a grid-cell average and we obtain

$$\langle \tau \rangle = \langle \overline{T} \rangle \vec{u} \tag{33}$$

for the estimate of drag at the coarse resolution model gridpoints. The velocity perturbation becomes unbounded (resonant) at the latitudes where $\omega = |f|$, but the drag itself vanishes smoothly as these latitudes are approached from the equator. We set the drag to zero poleward of the latitude where f equals the M_2 frequency.

Over length scales small enough for the background flow to appear steady ($\omega \ll |\vec{u}||\vec{k}|$) the nonlinearity parameter is the ratio of perturbation to background velocities $|\vec{u'}|/|\vec{u}|$, i.e., the Froude number

$$Fr = \frac{Nh}{|\vec{u}|}. \tag{34}$$

In the limit $\omega \gg |\vec{k}||\vec{u}|$, the nonlinearity parameter is the ratio of nonlinear advection to the time tendency term, i.e. $|\vec{k}||\vec{u'}|/\omega$. If we estimate $|\vec{k}||\vec{u'}|/\omega \approx \nabla \cdot \vec{u'}/\omega$ and use $\nabla \cdot \vec{u'} = \partial w/\partial z$ with the latter given by (27), we obtain $Fr(|\vec{u}||\vec{k}|/\omega)^2$ as our estimate of nonlinearity on large scales. The suggestion therefore is that nonlinearities are small on large scales. Flow splitting and nonlinear drag appear to be operating primarily when $|\vec{k}| \gg \omega/|\vec{u}|$, i.e., over topographic wavelengths less than about 400 m for M_2.

We now discuss the dimensional reasoning employed to include "nonpropagating" drag due to deflected flow and turbulence at the bottoms of large mountains. The dimensional reasoning we employ uses the Froude number $Nh/|\vec{u}|$ for steady flow (i.e., flow at small scales) as the nonlinearity parameter. We assume that the topography is characterized by well-defined features that can be binned into height ranges. If the mountain height exceeds a critical value h_c, the flow is blocked or deflected below a level $z = h - h_c$. For terrain features with heights less than h_c, the drag is entirely linear and propagating. For features with heights greater than h_c the drag includes both a propagating and a nonpropagating contribution. Internal waves are launched by the upper part of the mountain (over a vertical length scale h_c). In order to relate the width L of the "radiating" part of the mountain to elevation above the base, we introduce a power law with exponent δ:

$$L(z) = L_{\mathrm{base}}(1 - z/h)^{\delta}, \tag{35}$$

where L_{base} is the width of the mountain base. We take $\delta = 1$ in the utilization of our scheme, which corresponds to an assumption that mountains are triangular, but we note that our drag scheme is not overly sensitive to this parameter.

For radiating linear waves, the drag can be scaled as $\rho N V h^2/L$, where V is the magnitude of the background velocity and L is the mountain length scale. The scaling for nonpropagating drag is $\rho V^2(h - h_c)/L$. We set the nondimensional critical mountain height (or critical Froude number) $Fr_c = h_c(N/V)$.

The combination of these arguments is

$$D_{\mathrm{p}} = a_0 \min(1, (Fr_c/Fr)^{2-\delta}) Fr^2 (\overline{\rho} V^3 / N L_{\mathrm{base}}),$$

$$D_{\mathrm{np}} = a_1 (1 - \min(1, (Fr_c/Fr_0)^{1+\delta})) \times \frac{Fr_0}{1+\delta} (\overline{\rho} V^3 / N L_{\mathrm{base}}), \tag{36}$$

where D_{p} and D_{np} refer to the propagating and nonpropagating parts of the base flux, respectively, and a_0 and a_1 are constant drag coefficients.

To evaluate (36) over the mountains contained by a gridpoint in the coarse resolution model, we must have a relationship between h and L_{base}. We assume that

$$L/L_0 = (h/h_0)^{\gamma}, \tag{37}$$

where L_0, h_0, and γ are constants (we denote the Froude number corresponding to h_0 by Fr_0). We determine the latter empirically to be about 0.35, but again, our model is not overly sensitive to this parameter.

Integrating (36) under the stated assumptions about mountain height distributions yields

$$\langle D_{\text{p}} \rangle = a_0 Fr_0^{\gamma} \frac{H^{<}(2+\gamma) + H^{>}(\gamma-\delta)Fr_{\text{c}}^{2+\delta}}{H(2\gamma)} \frac{\overline{\rho} V^3}{NL_0},$$

$$\langle D_{\text{np}} \rangle = a_1 Fr_0^{\gamma} \frac{H^{>}(1+\gamma) - H^{>}(\gamma-\delta)Fr_{\text{c}}^{1+\delta}}{(1+\delta)H(2\gamma)} \times \frac{\overline{\rho} V^3}{NL_0}, \tag{38}$$

where brackets denote spatial averaging, $H(\alpha) = [Fr_{\max}^{\alpha} - Fr_{\min}^{\alpha}]/\alpha$, and the superscript on H means that Fr is replaced with either $Fr^{<} = \min(Fr, Fr_{\text{c}})$ or $Fr^{>} = \max(Fr, Fr_{\text{c}})$. The minimum and maximum values of the Froude number over the unresolved topography (denoted by subscripts) serve as limits of the area integration. If we let $Fr \to 0$ in $\langle D_{\text{p}} \rangle$ we obtain

$$D^* = a_0 \overline{\rho} NV \Omega^2 / L_0,$$

$$\Omega = Fr_0^{\gamma/2} \frac{V}{N} \left[\frac{H(2+\gamma)}{H(2\gamma)} \right]^{1/2}. \tag{39}$$

Refinements of the dimensional analyses above are presented in Garner (2004, submitted).

We allow the linear result (33) to determine the direction of drag, and modify our drag estimate by the dimensional analysis as

$$\langle \tau \rangle = \left(\frac{\langle D_{\text{p}} \rangle}{D^*} + \frac{\langle D_{\text{np}} \rangle}{D^*} \right) \tau^*, \tag{40}$$

where τ now refers to the final, corrected drag, and τ^* refers to the original linear result (33). The ratio of nonlinear to linear drag coefficients a_1/a_0 remains undetermined, as does the critical mountain height Fr_{c}. In the current paper we set the former to 10 and the latter to 1, consistent with optimal values determined in atmospheric simulations.

Even using the best currently available global topographic datasets, it is impossible to directly estimate topographic heights, and thus Froude numbers, at horizontal scales of 400 m or less. In order to take advantage of quantities such as χ that were calculated over resolved scales, we computed

$$Fr_{\text{proxy}} = Fr_{\text{largescale}} \varepsilon(\phi), \tag{41}$$

as a proxy for the Froude number, where

$$Fr_{\text{largescale}} = \frac{|\nabla^2 \chi_{\text{steady}}| L}{V}, \tag{42}$$

and

$$\varepsilon(\phi) = \frac{V\omega}{2L(\omega^2 - f^2)}. \tag{43}$$

The quantity $Fr_{\text{largescale}}$ scales like a Froude number Nh/V. The extrapolation factor $\varepsilon(\phi)$ accounts for the reduction in topographic heights from those associated with resolved topographic scales to those associated with the unresolved horizontal length scales of interest. A typical tidal velocity is denoted by V, while L denotes a typical horizontal scale in the topography. The factor $\omega/(\omega^2 - f^2)$ is introduced into $\varepsilon(\phi)$ in order to make the Froude number large as ω approaches f, so that most of the drag is nonlinear at the resonant latitudes for the tides. As in the linear analysis, we use the M_2 frequency to calculate the Froude numbers. Since we are interested in Froude numbers for the small-scale (steady) limit, the quantity χ_{steady} must be used, where χ_{steady} is the equivalent of χ in the steady flow case—χ_{steady} is calculated as in (31) but with N_* replaced by N (Garner, 2003; and submitted 2004). Maps of Fr_{proxy} (not shown) indicate that it is of order one in areas of rough topography, and we used it in the dimensional analysis as if it was the Froude number calculated on the small scales where flow-splitting takes place. In the future we may consider whether more reliable estimates of the Froude number at small scales can be made. Perhaps the extrapolation from large to small scales could utilize (37). With $a_1/a_0 = 10$, $(\langle D_{\text{p}} \rangle +$

$\langle D_{np} \rangle)/D^*$ does not vary by more than a factor of two over all Froude numbers. Thus alterations of the Froude number cannot change the drag by more than a factor of two. Such factors would be absorbed into our multiplicative factor.

References

Accad, Y., Pekeris, C.L., 1978. Solution of the tidal equations for the M_2 and S_2 tides in the world oceans from a knowledge of the tidal potential alone. Philosophical Transactions of the Royal Society of London A290, 235–266.

Andersen, O.B., Woodworth, P.L., Flather, R.A., 1995. Intercomparison of recent ocean tide models. Journal of Geophysical Research 100, 25261–25282.

Arbic, B.K., Flierl, G.R., 2004. Baroclinically unstable geostrophic turbulence in the limits of strong and weak bottom Ekman friction: application to mid-ocean eddies. Journal of Physical Oceanography 34, 2257–2273.

Armi, L., 1978. Some evidence for boundary mixing in the deep sea. Journal of Geophysical Research 83, 1971–1979.

Baines, P.G., 1973. The generation of internal tides by flat-bump topography. Deep-Sea Research 20, 179–205.

Baines, P.G., 1982. On internal tide generation models. Deep-Sea Research 29, 307–338.

Balmforth, N.J., Ierley, G.R., Young, W.R., 2002. Tidal conversion by subcritical topography. Journal of Physical Oceanography 32, 2900–2914.

Bell, T.H., 1975a. Lee waves in stratified flows with simple harmonic time dependence. Journal of Fluid Mechanics 67, 705–722.

Bell, T.H., 1975b. Topographically generated internal waves in the open ocean. Journal of Geophysical Research 80, 320–327.

Carrere, L., Lyard, F., 2003. Modeling the barotropic response of the global ocean to atmospheric wind and pressure forcing—comparisons with observations. Geophysical Research Letters 30, 1275 doi:10.1029/2002GL016473.

Cartwright, D.E., 1977. Oceanic tides. Reports on Progress in Physics 40, 665–708.

Cartwright, D.E., 1993. Theory of ocean tides with application to altimetry. In: Rummel, R., Sanso (Eds.), Satellite Altimetry in Geodesy and Oceanography. Springer, New York, pp. 99–141.

Cartwright, D.E., 1999. Tides: A Scientific History. Cambridge University Press, Cambridge, p. 292.

Cartwright, D.E., Edden, A.C., 1973. Corrected tables of tidal harmonics. Geophysical Journal of the Royal Astronomical Society 33, 253–264.

Cartwright, D.E., Ray, R.D., 1990. Oceanic tides from Geosat altimetry. Journal of Geophysical Research 95, 3069–3090.

Cartwright, D.E., Tayler, R.J., 1971. New computations of the tide-generating potential. Geophysical Journal of the Royal Astronomical Society 23, 45–74.

Cox, C.S., Sandstrom, H., 1962. Coupling of surface and internal waves in water of variable depth. Journal of the Oceanographic Society of Japan 20th Anniversary Volume 499–513.

Cummins, P.F., Oey, L.Y., 1997. Simulation of barotropic and baroclinic tides off Northern British Columbia. Journal of Physical Oceanography 27, 762–781.

Cummins, P.F., Cherniawski, Y., Foreman, M.G.G., 2001 North Pacific internal tides from the Aleutian Ridge: altimeter observations and modeling. Journal of Marine Research 59, 167–191.

Desai, S., Wahr, J.M., 1995. Empirical ocean tide models estimated from TOPEX/POSEIDON altimetry. Journal of Geophysical Research 100, 25205–25228.

Desai, S.D., Wahr, J.M., Chao, Y., 1997. Error analysis of empirical ocean tide models estimated from TOPEX/POSEIDON altimetry. Journal of Geophysical Research 102, 25157–25172.

Doodson, A.T., 1921. Harmonic development of the tide-generating potential. Proceedings of the Royal Society of London A100, 305–329.

Dushaw, B.D., Cornuelle, B.D., Worcester, P.F., Howe, B.M., Luther, D.S., 1995. Barotropic and baroclinic tides in the central North Pacific Ocean determined from long-range reciprocal acoustic transmissions. Journal of Physical Oceanography 25, 631–647.

Egbert, G.D., Erofeeva, S.Y., 2002. Efficient inverse modeling of barotropic ocean tides. Journal of Atmospheric and Oceanic Technology 19, 183–204.

Egbert, G.D., Ray, R.D., 2000. Significant dissipation of tidal energy in the deep ocean inferred from satellite altimeter data. Nature 405, 775–778.

Egbert, G.D., Ray, R.D., 2001. Estimates of M_2 tidal energy dissipation from TOPEX/Poseidon altimeter data. Journal of Geophysical Research 106, 22475–22502.

Egbert, R.D., Ray, R.D., 2003a. Semi-diurnal and diurnal tidal dissipation from TOPEX/Poseidon altimetry. Geophysical Research Letters 30, 1907 doi:10.1029/2003GL017676.

Egbert, G.D., Ray, R.D., 2003b. Deviation of long-period tides from equilibrium: kinematics and geostrophy. Journal of Physical Oceanography 33, 822–840.

Egbert, G.D., Bennett, A.F., Foreman, M.G.G., 1994. TOPEX/POSEIDON tides estimated using a global inverse model. Journal of Geophysical Research 99, 24821–24852.

Egbert, G.D., Ray, R.D., Bills, B.G., 2004. Numerical modeling of the global semidiurnal tide in the present day and in the last glacial maximum. Journal of Geophysical Research 109, C03003 doi:10.1029/2003JC001973.

Farrell, W.E., 1972. Deformation of the earth by surface loads. Reviews of Geophysics and Space Physics 10, 761–797.

Flierl, G.R., 1978. Models of vertical structure and the calibration of two-layer models. Dynamics of Atmospheres and Oceans 2, 341–381.

Foreman, M.G.G., 1977. Manual for tidal heights analysis and prediction. Pacific Marine Science Report 77-10, Institute of Ocean Sciences, 101pp.

Francis, O., Mazzega, P., 1990. Global charts of ocean tide loading effects. Journal of Geophysical Research 95, 11411–11424.

Garner, S.T., 2003. A topographic drag closure with analytical base flux. Paper P3.1, 14th Conference on Atmospheric and Oceanic Fluid Dynamics, American Meteorological Society.

Gordeev, R.G., Kagan, B.A., Polyakov, E.V., 1977. The effects of loading and self-attraction on global ocean tides: the model and the results of a numerical experiment. Journal of Physical Oceanography 7, 161–170.

Hallberg, R., Rhines, P., 1996. Buoyancy-driven circulation in an ocean basin with isopycnals intersecting the sloping boundary. Journal of Physical Oceanography 26, 914–940.

Hendershott, M.C., 1972. The effects of solid earth deformation on global ocean tides. Geophysical Journal of the Royal Astronomical Society 29, 389–402.

Hendershott, M.C., 1977. Numerical modeling of ocean tides. In: Goldberg, E.D., McCave, I.N., O'Brien, J.J., Steele, J.H. (Eds.), The Sea, vol. 6. Wiley, New York, pp. 47–95.

Hendershott, M.C., 1981. Long waves and ocean tides. In: Warren, B.A., Wunsch, C. (Eds.), Evolution of Physical Oceanography. MIT Press, Cambridge, pp. 292–341.

Hendry, R.M., 1977. Observations of the semidiurnal internal tide in the western North Atlantic Ocean. Philosophical Transactions of the Royal Society of London A286, 1–24.

Hibiya, T., 1986. Generation mechanism of internal waves by tidal flow over a sill. Journal of Geophysical Research 91, 7696–7708.

Hirose, N., Fukimori, I., Zlotnicki, V., Ponte, R.M., 2001. Modeling the high-frequency barotropic response of the ocean to atmospheric disturbances: sensitivity to forcing, topography, and friction. Journal of Geophysical Research 106, 30987–30995.

Holloway, P.E., 1996. A numerical model of internal tides with application to the Australia North West Shelf. Journal of Physical Oceanography 26, 21–37.

Holloway, P.E., Merrifield, M.A., 1999. Internal tide generation by seamounts, ridges, and islands. Journal of Geophysical Research 104, 25937–25951.

Holton, J.R., 1972. An Introduction to Dynamic Meteorology. Academic Press, New York, 319pp.

IOC, IHO, BODC, 2003. Centenary edition of the GEBCO digital atlas. Published on CD-ROM on behalf of the Intergovernmental Oceanographic Commission and the International Hydrographic Organization as part of the General Bathymetric Chart of the Oceans, British Oceanographic Data Centre, Liverpool.

Jakobssen, M., Chervis, N.Z., Woodward, J., Macnab, R., Coakley, B., 2000. New grid of Arctic bathymetry aids scientists and mapmakers. Eos Transactions, American Geophysical Union 81, 89,93,96.

Jayne, S.R., St. Laurent, L.C., 2001. Parameterizing tidal dissipation over rough topography. Geophysical Research Letters 28, 811–814.

Jeffreys, H., 1920. Tidal friction in shallow seas. Philosophical Transactions of the Royal Society of London A221, 239–264.

Kagan, B.A., Sundermann, J., 1996. Dissipation of tidal energy, paleotides, and evolution of the Earth–Moon system. Advances in Geophysics 38, 179–266.

Kang, S.K., Foreman, M.G.G., Crawford, W.R., Cherniawsky, J.Y., 2000. Numerical modeling of internal tide generation along the Hawaiian Ridge. Journal of Physical Oceanography 30, 1083–1098.

Kantha, L.H., 1998. Tides: a modern prospective. Marine Geodesy 21, 275–298.

Kantha, L.H., Tierney, C.G., 1997. Global baroclinic tides. Progress in Oceanography 40, 163–178.

Khatiwala, S., 2003. Generation of internal tides in an ocean of finite depth: analytical and numerical calculations. Deep-Sea Research I 50, 3–21.

Kunze, E., Toole, J.M., 1997. Tidally driven vorticity, diurnal shear, and turbulence atop Fieberling Seamount. Journal of Physical Oceanography 27, 2663–2693.

Ledwell, J.R., Montgomery, E.T., Polzin, K.L., St. Laurent, L.C., Schmitt, R.W., Toole, J.M., 2000. Evidence for enhanced mixing over rough topography in the abyssal ocean. Nature 403, 179–182.

Legg, S., 2004. Internal tides generated on a corrugated continental slope. Part I: Cross-slope barotropic forcing. Journal of Physical Oceanography 34, 156–173.

Le Provost, C., Lyard, F., 1997. Energetics of the M_2 barotropic ocean tides: an estimate of bottom friction dissipation from a hydrodynamic model. Progress in Oceanography 40, 37–52.

Le Provost, C., Genco, M.L., Lyard, F., Vincent, P., Canceil, P., 1994. Spectroscopy of the world ocean tides from a finite element hydrodynamic model. Journal of Geophysical Research 99, 24777–24797.

Le Provost, C., Bennett, A.F., Cartwright, D.E., 1995. Ocean tides for and from TOPEX/POSEIDON. Science 267, 639–642.

Levitus, S., Boyer, T.P., Conkright, M.E., O'Brien, T., Antonov, J., Stephens, C., Stathoplos, L., Johnson, D., Gelfeld, R., 1998. World Ocean Database 1998, vol. 1: Introduction. NOAA Atlas NESDIS 18, U.S. Government Printing Office, Washington, D.C., 346pp.

Llewellyn Smith, S.G., Young, W.R., 2002. Conversion of the barotropic tide. Journal of Physical Oceanography 32, 1554–1566.

Llewellyn Smith, S.G., Young, W.R., 2003. Tidal conversion at a very steep ridge. Journal of Fluid Mechanics 495, 175–191.

Lott, F., Miller, M.J., 1997. A new subgrid-scale orographic drag parameterization: its formulation and testing. Quarterly Journal of the Royal Meteorological Society 123, 101–127.

Lueck, R.G., Mudge, T.D., 1997. Topographically induced mixing around a shallow seamount. Science 276, 1831–1833.

Maas, L.R., Doelman, A., 2002. Chaotic tides. Journal of Physical Oceanography 32, 870–890.

Marchuk, G.I., Kagan, B.A., 1984. Ocean Tides. Pergamon Press, Oxford, 292pp.

Merrifield, M.A., Holloway, P.E., Shaun Johnston, T.M., 2001. The generation of internal tides at the Hawaiian Ridge. Geophysical Research Letters 28, 559–562.

Miller, G.R., 1966. The flux of tidal energy out of the deep ocean. Journal of Geophysical Research 71, 2485–2489.

Miranda, M.A., James, I.N., 1992. Non-linear three-dimensional effects on gravity-wave drag: splitting flow and breaking waves. Quarterly Journal of the Royal Meteorological Society 118, 1057–1081.

Morozof, E.G., 1995. Semidiurnal internal wave global field. Deep-Sea Research 42, 135–148.

Munk, W., 1997. Once again: once again—tidal friction. Progress in Oceanography 40, 7–35.

Munk, W.H., MacDonald, G.J.F., 1960. The Rotation of the Earth. Cambridge University Press, London, 323pp.

Munk, W., Wunsch, C., 1998. Abyssal recipes II: energetics of tidal and wind mixing. Deep-Sea Research I 45, 1977–2010.

Niwa, Y., Hibiya, T., 2001. Numerical study of the spatial distribution of the M_2 internal tide in the Pacific Ocean. Journal of Geophysical Research 106, 22441–22449.

Padman, L., Erofeeva, S., 2004. A barotropic inverse tidal model for the Arctic Ocean. Geophysical Research Letters 31, L02303 doi:10.1029/2003GL019003.

Padman, L., Fricker, H.A., Coleman, R., Howard, S., Erofeeva, S., 2002. A new tidal model for the Antarctic ice shelves and seas. Annals of Glaciology 34, 247–254.

Padman, L., Erofeeva, S., Joughin, I., 2003. Tides of the Ross Sea and Ross Ice Shelf cavity. Antarctic Science 15, 31–40.

Parke, M.E., Hendershott, M.C., 1980. M2, S2, K1 models of the global ocean tide on an elastic earth. Marine Geodesy 3, 379–408.

Pawlowicz, R., Beardsley, B., Lentz, S., 2002. Classical tidal harmonic analysis including error estimates in MATLAB using T-TIDE. Computers and Geosciences 28, 929–937.

Pierrehumbert, R.T., 1987. An essay on the parameterization of orographic gravity wave drag. Proceedings seminar/workshop on observations, theory and modeling of orographic effects, vol. 1, Shinfield Park, Reading. UK, ECMWF, pp. 251 -282.

Platzman, G.W., Curtis, G.A., Hansen, K.S., Slater, R.D., 1981. Normal modes of the world ocean. Part II: Description of modes in the period range 8–80 h. Journal of Physical Oceanography 11, 579–603.

Polzin, K., 2004. Idealized solutions for the energy balance of the finescale internal wave field. Journal of Physical Oceanography 34, 231–246.

Polzin, K.L., Toole, J.M., Ledwell, J.R., Schmitt, R.W., 1997. Spatial variability of turbulent mixing in the abyssal ocean. Science 276, 93–96.

Pugh, D.T., 1987. Tides, Surges and Mean Sea-level: A Handbook for Scientists and Engineers. Wiley, Chichester, 472pp.

Ray, R.D., 1994. Tidal energy dissipation: observations from astronomy, geodesy, and oceanography. In: Majumdar, S., et al. (Eds.), Pennsylvania Academy of Sciences, Easton, PA, pp. 171–185.

Ray, R.D., 1998. Ocean self-attraction and loading in numerical tidal models. Marine Geodesy 21, 181–192.

Ray, R.D., 1999. A global ocean tide model from TOPEX/POSEIDON altimetry: GOT99.2. National Aeronautics and Space Administration Technical Memorandum, NASA/TM-1999-209478, 58pp.

Ray, R.D., Mitchum, G.T., 1996. Surface manifestation of internal tides generated near Hawaii. Geophysical Research Letters 23, 2101–2104.

Ray, R.D., Mitchum, G.T., 1997. Surface manifestation of internal tides in the deep ocean: observations from altimetry and tide gauges. Progress in Oceanography 40, 135–162.

Roosbeek, F., 1996. RATGP95: a harmonic development of the tide-generating potential using an analytical method. Geophysical Journal International 126, 197–204.

Rudnick, D.L., Boyd, T.J., Brainard, R.E., Carter, G.S., Egbert, G.D., Gregg, M.C., Holloway, P.E., Klymak, J.M., Kunze, E., Lee, C.M., Levine, M.D., Luther, D.S., Martin, J.P., Merrifield, M.A., Moum, J.N., Nash, J.D., Pinkel, R., Rainville, L., Sanford, T.B., 2003. From tides to mixing along the Hawaiian Ridge. Science 301, 355–357.

Schwiderski, E.W., 1980. On charting global ocean tides. Reviews of Geophysics and Space Physics 18, 243–268.

Shum, C.K., Woodworth, P.L., Andersen, O.B., Egbert, G.D., Francis, O., King, C., Klosko, S.M., Le Provost, C., Li, X., Molines, J.-M., Parke, M.E., Ray, R.D., Schlax, M.G., Stammer, D., Tierney, C.C., Vincent, P., Wunsch, C.I., 1997. Accuracy assessment of recent ocean tide models. Journal of Geophysical Research 102, 25173–25194.

Simmons, H.L., Hallberg, R.W., Arbic, B.K., 2004. Internal wave generation in a global baroclinic tide model. Deep-Sea Research II, this issue [doi:10.1016/j.dsr2.2004.09.015].

Sjoberg, B., Stigebrandt, A., 1992. Computations of the geographical distribution of the energy flux to mixing processes via internal tides and the associated vertical circulation in the ocean. Deep-Sea Research 39, 269–291.

Smith, W.H.F., Sandwell, D.T., 1997. Global sea floor topography from satellite altimetry and ship depth soundings. Science 277, 1956–1962.

St. Laurent, L.C., Garrett, C., 2002. The role of internal tides in mixing the deep ocean. Journal of Physical Oceanography 32, 2882–2899.

St. Laurent, L.C., Stringer, S., Garrett, C., Perrault-Joncas, D., 2003. The generation of internal tides at abrupt topography. Deep-Sea Research 50, 987–1003.

Taylor, G.I., 1919. Tidal friction in the Irish Sea. Philosophical Transactions of the Royal Society of London A220, 1–93.

Thurnherr, A.M., Richards, K.J., 2001. Hydrography and high-temperature heat flux of the Rainbow hydrothermal site (36°14′N, Mid-Atlantic Ridge). Journal of Geophysical Research 106, 9411–9426.

Thurnherr, A.M., Speer, K.G., 2003. Boundary mixing and topographic blocking on the Mid-Atlantic Ridge in the

South Atlantic. Journal of Physical Oceanography 33, 848–862.

Thurnherr, A.M., Richards, K.J., German, C.R., Lane-Serff, G.F., Speer, K.G., 2002. Flow and mixing in the rift valley of the Mid-Atlantic Ridge. Journal of Physical Oceanography 32, 1763–1778.

Tierney, C.G., Kantha, L.H., Born, G.H., 2000. Shallow and deep water global ocean tides from altimetry and numerical modeling. Journal of Geophysical Research 105, 11259–11277.

Wahr, J.M., 1981. Body tides on an elliptical, rotating, elastic and oceanless Earth. Geophysical Journal of the Royal Astronomical Society 64, 677–703.

Wahr, J.M., Sasao, T., 1981. A diurnal resonance in the ocean tide and in the Earth's load response due to the resonant free "core nutation". Geophysical Journal of the Royal Astronomical Society 64, 747–765.

Wunsch, C., 1975. Internal tides in the ocean. Reviews of Geophysics and Space Physics 13, 167–182.

Deep-Sea Research Part II – Instructions to Authors

A more detailed set of instructions can be found in the journal's home page at http://www.elsevier.com/locate/dsr2

1. Submission of Papers

Authors are requested to submit their original manuscript and figures with two copies to the Guest Editor of their Special Issue. Authors should ensure that they submit manuscripts and meet any additional requirements in line with deadlines set by the Guest Editor(s) to ensure that the entire issue can be published in a timely fashion. In most instances, revised manuscripts must be resubmitted within six months of receipt by the author. Failure to do so may lead to additional review, in some instances, rejection of the paper. Submission of a paper implies that it has not been published previously, that it is not under consideration for publication elsewhere, and that if accepted it will not be published elsewhere in the same form, in English or in any other language, without the written consent of the publisher.

2. Manuscript Preparation

General: Manuscripts must be typewritten, double-spaced with wide margins on one side of white paper. Good quality printouts with a font size of 12 or 10 pt are required. The corresponding author should be identified (include a Fax number and E-mail address). Full postal addresses must be given for all co-authors. Authors should consult a recent issue of the journal for style if possible. An electronic copy of the paper should accompany the final version. The Editors reserve the right to adjust style to certain standards of uniformity. Authors should retain a copy of their manuscript since we cannot accept responsibility for damage or loss of papers. Original manuscripts are discarded one month after publication unless the Publisher is asked to return original material after use.

Paper Length: Authors should try to confine the text to no more than 25–30 pages, with a total manuscript length (including figures and tables) of 45–50 pages.

Abstracts: The essential contents of each paper should be recapitulated briefly in an abstract of no more than 500 words at the head of the paper.

Keywords: Please provide up to six subject keywords, plus regional index terms. At least four of the subject keywords should be selected from the *Aquatic Science & Fisheries Thesaurus*. An eletronic version of the Thesaurus can be found at: http://www.csa.com/ helpV3/ ab.html. The Regional terms should be provided as a hierarchical string, eg; USA, California, Monterey Bay.

Text: Follow this order when typing manuscripts: Title, Authors, Affiliations, Abstract, Keywords, Main text, Acknowledgements, Appendix, References, Vitae, Figure Captions and then Tables. The corresponding author should be identified with an asterisk and footnote. All other footnotes (except for table footnotes) should be identified with superscript Arabic numbers.

Units: In general, metric units should be used, preferably the Systeme International (SI). If other units are used (fathoms, F, knots, etc), the metric equivalents must be given in parentheses ( ) or by presenting the dimensionally correct conversion factors as a footnote.

References: All publications cited in the text should be presented in a list of references following the text of the manuscript. In the text refer to the author's name (without initials) and year of publication (e.g. "Since Peterson (1993) has shown that ..." or "This is in agreement with results obtained later (Kramer, 1994)". For three or more authors use the first author followed by "et al.", in the text. The list of references should be arranged alphabetically by authors' names. The manuscript should be carefully checked to ensure that the spelling of authors' names and dates are exactly the same in the text as in the reference list. The title of the journal should not be abbreviated; titles of journals in Russian language should be transliterated and not translated. References should be given in the following form:

Gooday, A.J., Bett, B.J., Shires, R., Lambshead, P.J.D., 1998. Deep-sea benthic foraminiferal species diversity in the NE Atlantic and NW Arabian Sea: a synthesis. Deep-Sea Research Part II 45 (1–3), 165–201.

Hessler, R.R., Wilson, G.D.F., 1983. The origin and biogeography of malacostracan crustaceans in the deep sea. In: Sims, R.W., Price, J.H., Whalley, P.E.S. (Eds.), Evolution, time and space: the emergence of the biosphere. Academic Press. London, pp. 227–254. Thiel, H., 1983. Benthos size classification. Deep-Sea Newsletter 20, 8–11.

References to unpublished literature should be set out in the format detailed below. The following information needs to be provided: Names and initials of all authors., year. Title of work. All other relevant information needed to identify items (e.g., technical report, Ph.D. Thesis, institute, current status, i.e. in press/unpublished, etc).

Marstakas, N., 1990. Relationships of morphological and physicochemical properties of vertisols under Greek climate conditions. Ph.D. Thesis, Agricultural University Athens, unpublished.

Examples of appropriate citation formats for non-text and electronic-only information can be found in the Guide for Authors at the Deep-Sea Research Part II website. However, it is requested that a web site address, or listserver message is given as a reference only where the information is unavailable in a more permanent form. If such sources are given, then please give as complete information as is possible. All references to such material in the text of the article must be supported by a full entry in the bibliography.

Illustrations: All illustrations should be provided in camera-ready form, suitable for reproduction (which may include reduction) without retouching. Photographs, charts and diagrams are all to be referred to as "Figure(s)" and should be numbered consecutively in the order to which they are referred. They should accompany the manuscript, but should not be included within the text. All illustrations should be clearly marked on the back with the figure number and the author's name. All figures are to have a caption. Captions should be supplied on a separate sheet.

Line drawings: Good quality printouts on white paper produced in black ink are required. All lettering, graph lines and points c graphs should be sufficiently large and bold to permit reproduction when the diagram has been reduced to a size suitable for inclusio in the journal. Dye-line prints or photocopies are not suitable for reproduction. Do not use any type of shading on computer-generate illustrations.
Photographs: Original photographs must be supplied as they are to be reproduced (e.g. black and white or colour). If necessary, a sca should be marked on the photograph. Please note that photocopies of photographs are not acceptable. Half-tone illustrations shoul be kept to a minimum.
Fold-charts: The cost of fold-charts or diagrams must be borne by the author or his institution. *Colour:* Authors will be charged fc colour reproduction as follows: €350 (or US$350) for the first page, €175 (or US$175) for subsequent pages.
Submit colour illustrations as original photographs, high-quality computer prints or transparencies, close to the size expected i publication, or as 35 mm slides. Polaroid colour prints are not suitable. If, together with your accepted article, you submit usabl colour figures then Elsevier will ensure, at no additional charge, that these figures will appear in colour on the web (e.g., ScienceDirec and other sites) regardless of whether or not these illustrations are reproduced in colour in the printed version. For colour reproductio in print, you will receive information regarding the costs from Elsevier after receipt of your accepted article. For further information o the preparation of electronic artwork, please see *Artwork Instructions*. Please note: Because of technical complications which can aris by converting colour figures to 'grey scale' (for the printed version should you not opt for colour in print) please submit in additio usable black and white prints corresponding to all the colour illustrations.
Tables: Tables should be numbered consecutively and given a suitable caption and each table typed on a separate sheet. Footnotes t tables should be typed below the table and should be referred to by superscript lowercase letters. No vertical rules should be used Tables should not duplicate results presented elsewhere in the manuscript, (e.g. in graphs).

Authors in Japan kindly note: Upon request Elsevier Japan will provide a list of people who can check and improve the English of ar article (*before submission*). Please contact our Tokyo office: Elsevier Japan K.K., 1-9-15 Higashi Azabu, Minato-ku, Tokyo 106-0044 Japan; tel.: +81-3-5561-5032; fax: +81-3-5561-5045; e-mail: jp.info@elsevier.com

3. Electronic Submission

Authors should submit an electronic copy of their paper with the final version of the manuscript. The electronic copy should match the hardcopy exactly. For authors using LaTex, the document style files, as well as the instructions "Preparing articles with LaTex" in the form of a dvi file, can be obtained free of charge from any host on the Comprehensive Tex Archive Network (CTAN) using anonymous ftp. The primary CTAN hosts are ftp.dante.de and ftp.tex.ac.uk. The Elsevier macros are in /pub/tex/macros/latex/contrib/supported/ elsevier. These can also be downloaded from the Internet site http://www.tex.ac.uk. Always keep a back up copy of the electronic file for reference and safety. Full details of electronic submission and other formats can be obtained from http://authors.elsevier.com or from Author Support at Elsevier.

4. Proofs

Proofs will be sent to the author (first named author if no corresponding author is identified of multi-authored papers) and should be returned within 48 hours of receipt. Corrections should be restricted to typesetting errors; any others may be charged to the author. Any queries should be answered in full. Please note that authors are urged to check their proofs carefully before return, since the inclusion of late corrections cannot be guaranteed. If authors are expecting to be away from their office when the proofs arrive, arrangements should be made with colleagues for the proofs to be checked in the authors absence. Proofs are to be returned to the Log-in Department, Elsevier, Stover Court, Bampfylde Street, Exeter, Devon EX1 2AH, UK.

5. Offprints

Twenty-five offprints and a copy of the issue will be supplied free of charge to each corresponding author of a paper. Additional offprints can be ordered at a specially reduced rate using the order form sent to the corresponding author after the manuscript has been accepted. Orders for reprints will incur a 50% surcharge.

6. Copyright

All authors must sign the "Transfer of Copyright" agreement before the article can be published. This transfer agreement enables Elsevier Ltd to protect the copyrighted material for the authors, but does not relinquish the author's proprietary rights. The copyright transfer covers the exclusive rights to reproduce and distribute the article, including reprints, photographic reproductions, microfilm or any other reproductions of similar nature and translations. Includes the right to adapt the article for use in conjunction with computer systems and programs, including reproduction or publication in machine-readable form and incorporation in retrieval systems. Authors are responsible for obtaining from the copyright holder permission to reproduce any figures for which copyright exists.

7. Author Enquiries

For enquiries relating to the submission of articles (including electronic submission), the preparation of electronic artwork, the status of accepted articles, author Frequently Asked Questions and any other enquiries relating to Elsevier Science, please consult http:// authors.elsevier.com

Contact details for questions arising after acceptance of an article, especially those relating to proofs, are provided when an article is accepted for publication.

Authors can also keep a track on the progress of their accepted article, and setup e-mail alerts informing them of changes to their manuscripts status, by using the "Track a paper" feature of Elsevier's Author gateway (http://authors.elsevier.com).

(*Continued from outside back cover*)

H.L. Simmons, R.W. Hallberg and B.K. Arbic	3043	Internal wave generation in a global baroclinic tide model
B.K. Arbic, S.T. Garner, R.W. Hallberg and H.L. Simmons	3069	The accuracy of surface elevations in forward global barotropic and baroclinic tide models

CONTENTS direct

This journal is part of **ContentsDirect**, the ***free*** alerting service which sends tables of contents by e-mail for Elsevier books and journals. You can register for **ContentsDirect** online at: **www.elsevier.com/locate/contentsdirect**

Indexed/abstracted in: *Curr Cont/Agri Bio Env Sci. Cirr Cont/Phys. Chem. Earth Sci, Curr Cont Sci Cit Ind, Res Alert, Global Change, Curr Aware BioSci, Geobase, Ocean Lit Rev, Awuat Sci. Fish Abstr, Meteor, Geoastrophys Abstr*

ISSN 0967–0645

ELSEVIER

DEEP-SEA RESEARCH PART II

Topical Studies in Oceanography

Small and mesoscale processes and their impact on the large scale

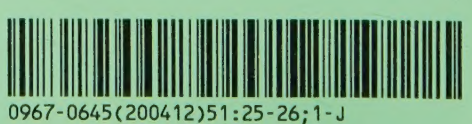
Volume 51, Numbers 25–26, 2004

CONTENTS

H. van Haren, L. St. Laurent and D. Marshall	2881	Preface
H. van Haren, L. St. Laurent and D. Marshall	2883	Small and mesoscale processes and their impact on the large scale: an introduction
B.R. Sutherland, M.R. Flynn and K. Dohan	2889	Internal wave excitation from a collapsing mixed region
R.H.J. Grimshaw and K.R. Khusnutdinova	2905	Internal waves in a three-layer bubbly waveguide
P. Bouruet-Aubertot, J. Sommeria, B. Le Cann and C.R. Koudella	2919	Intermittency of vertical density gradients at finescale and link with mixing processes
P. Hosegood and H. van Haren	2943	Near-bed solibores over the continental slope in the Faeroe-Shetland Channel
W.K. Dewar and D.A. Leonov	2973	Variability on steep, confined topography
T. Gerkema, F.-P.A. Lam and L.R.M. Maas	2995	Internal tides in the Bay of Biscay: conversion rates and seasonal effects
J. Isern-Fontanet, J. Font, E. García-Ladona, M. Emelianov, C. Millot and I. Taupier-Letage	3009	Spatial structure of anticyclonic eddies in the Algerian basin (Mediterranean Sea) analyzed using the Okubo–Weiss parameter
L.C. St. Laurent and J.D. Nash	3029	An examination of the radiative and dissipative properties of deep ocean internal tides

(*Continued on inside back cover*)

Visit our home page:
http://www.elsevier.com/locate/dsr2

Printed in Great Britain by Alden Press Ltd., Osney Mead, Oxford OX2 OEF

Also available on
SCIENCE DIRECT®
www.sciencedirect.com